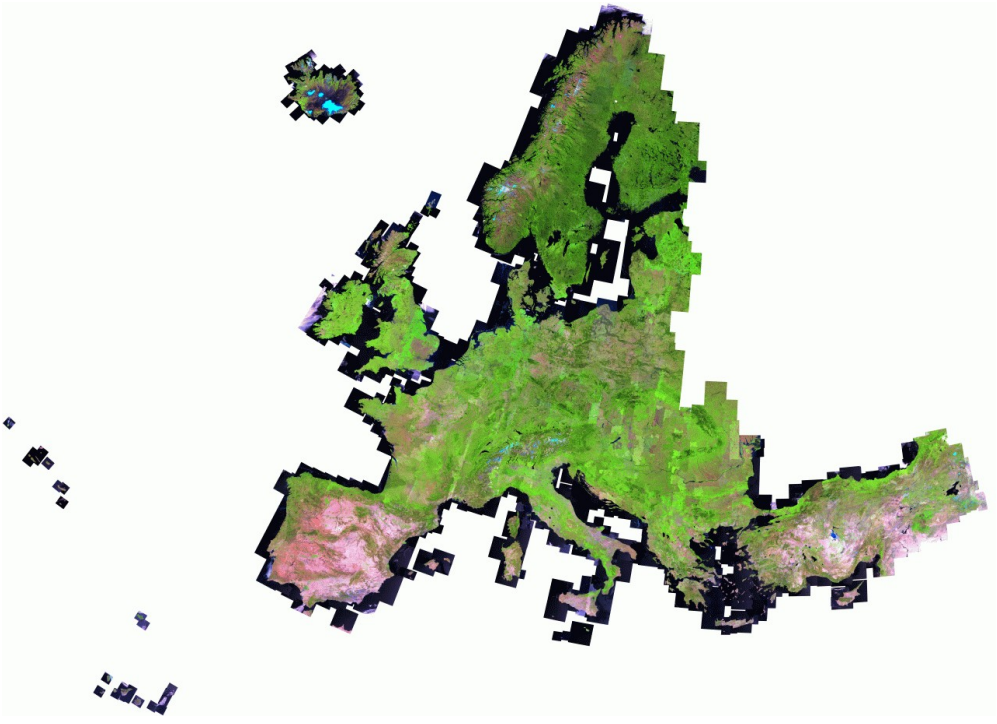
JRC Scientific and Technical Reports

The IMAGE-2006 Mosaic Project

Edited by Pierre Soille



JRC EUROPEAN COMMISSION

ies Institute for Environment and Sustainability

EUR 23755 EN - 2008

The mission of the Institute for Environment and Sustainability is to provide scientific-technical support to the European Union's Policies for the protection and sustainable development of the European and global environment.

European Commission
Joint Research Centre
Institute for Environment and Sustainability

Contact information
Address: via E. Fermi, 2749, I-21027 Ispra (Italy)
E-mail: Pierre.Soille@jrc.ec.europa.eu
Tel.: int+39-0332 785 068
Fax: int+39-0332 786 325

http://ies.jrc.ec.europa.eu/
http://www.jrc.ec.europa.eu/

JRC 49569

EUR 23755 EN
ISBN 978-92-79-20400-5
ISSN 1831-9424
doi:10.2788/25572

Luxembourg: Publications Office of the European Union, 2011

Printed in Italy

JRC
EUROPEAN COMMISSION

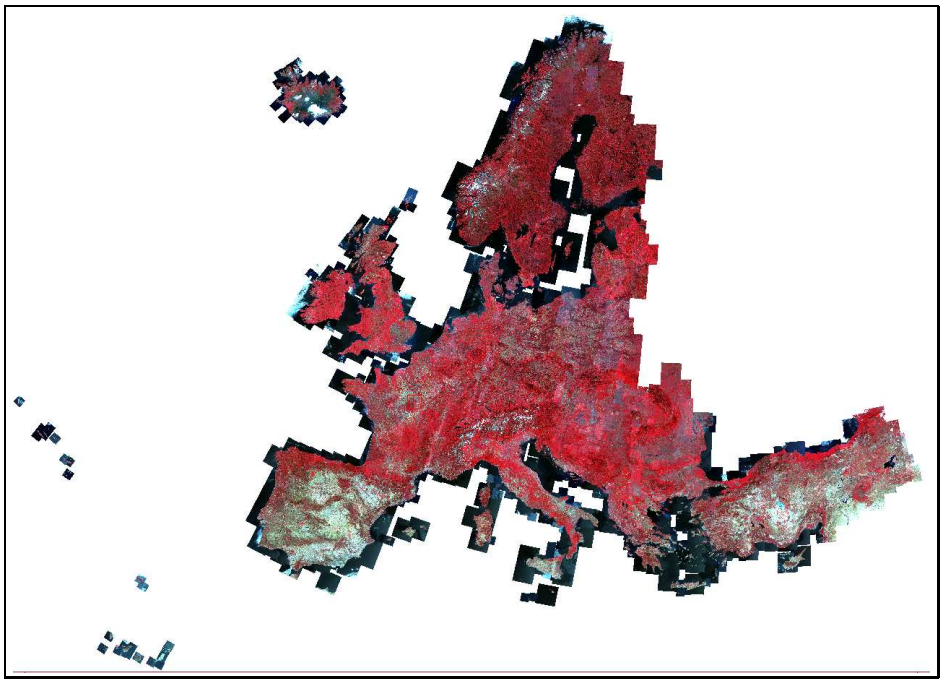
ies
Institute for Environment and Sustainability

The IMAGE-2006 Mosaic Project

Edited by Pierre Soille

Spatial Data Infrastructures Unit
Institute for Environment and Sustainability
Joint Research Centre, European Commission

December 23, 2008



Contents

Abstract

This report details the data ingestion, data organisation, and main processing steps adopted for generating the IMAGE2006 mosaic products. Chapter 1 describes the received data and the way it has been ingested and organised. Chapter 2 presents a detailed analysis of the footprints of the received imagery. Chapter 3 details the procedure used for cloud detection. Relative geometric and radiometric accuracy is studied in Chap. 4. The generation of a cloud mask is indeed fundamental for the generation of a mosaic minimising cloud cover. The mosaicing methodology is presented in Chap. 5. The produced mosaics are described in Chap. 6.

Chapter 1

Data ingestion and organisation

Pierre Soille and Conrad Bielski

Spatial Data Infrastructure Unit
Institute for Environment and Sustainability
DG Joint Research Centre, European Commission

1.1 Received data

The JRC received the data directly from DLR on portable hard disks on the 28th of February 2008 with a series of subsequent updates through FTP. The last update, on the 15th of April 2008, corrected for all the imagery of the Czech Republic because it did not follow the European reference grid specification [2]. The received data after this last update is referred to as VER1-0. A technical report describing procedures, methodologies, activities, and final results including statistics to document the orthoimage product generation project has been made available to JRC on the 14th of July 2008 [27].

The received data is organised on a per country basis with two coverages containing each imagery in European and national projections. The countries are referred to by the following country codes using lower case letters: al at ba be bg ch cs cy cz de dk ee es fi fr gb gr hr hu ie is it li lt lu lv mc me mt ni nl no pl pt ro se si sk tr. Note that there are 39 codes instead of 38 (actual number of participating countries) because the usual code for United Kingdom (UK) was divided into Great Britain (gb) and Northern Ireland (ni). This distinction originates from a special request from United Kingdom to process Northern Ireland in the Irish projection [27]. Note also that the code used for the Former Yugoslav Republic of Macedonia (FYROM) is mc and not the normalised ISO code MK [6]. Similarly, the code used for the republic of Serbia (cs) does not match the ISO code RS (cs corresponds to the code used for the former

```
2006_DLR/VER1-0/
|---al
|-----Coverage_1
|-------European
|-------National
|-------Original_IRSP6
|-------Original_Spot
|-----Coverage_2
|-------European
|-------National
|-------Original_IRSP6
|-------Original_Spot
|-----GCP
```

Figure 1.1: Directory tree structure as received from DLR: example for Albania (country code al). The GCPs were received separately and were added to each country directory.

union of Serbia and Montenegro). Figure 1.1 illustrates how the received data is organised for the case of Albania (country code al). Only the European and National directories are populated with imagery data. This data is encapsulated in an archive file (zip format), one for each image of the considered country and coverage. Figure 1.2 lists the content of sample IRS and SPOT archive files as stored in the European sub-directories.

The `imagery.bil` file contains the orthorectified imagery in either the ETRS-LAEA projection [3] following the European grid specification [2] (`European` directory) or the national projection (`National` directory). The `geolayer.bil` refers to the geolocation file. This two band file indicates the x-y coordinates of each pixel of the non-orthorectified (raw) imagery in the target projection (either European or National). Unfortunately, it has little value at the moment because the raw imagery was not included in the `Original_IRSP6` and `Original_Spot` directories. For IRS-P6, comprehensive metadata information related to the raw imagery is stored in an ASCII file in the Super Structure Data (SSD) format [10] while for SPOT this information is stored in an XML ASCII file in the Digital Image MAP (DIMAP) format [4]. Further to image data, the Ground Control Point (GCP) image chip database [27] downloaded from DLR by ftp on the 21st of February 2008 is stored under the directory `GCP` under each country directory. This database contains 61,053 chips of size 101×101 pixels stored in `png` format. The name of each chip is in the form `Easting_Northing_EllipsoidHeight.png`.

All received data described in this section is available to the IES SDI and LMNH units at the following location: `netapp2:/geodata/EUIM/2006_DLR`. The received imagery in the ETRS-LAEA projection is summarised in table 1.1 for each coverage, for each sensor, and for both coverages. The number between parenthesis indicates the number of unique source images. Indeed, the imagery for the European coverage has been delivered and processed on a country basis so that a scene overlapping two or more countries may have been orthorectified and delivered more than once with different ground control points and/or reference data. It follows that duplicate scenes are not necessarily identical (in

```
Directory:  2006_DLR/VER1-0/cz/Coverage_1/European/
Archive File:
IR06_LI3_ORT_1O_20050922T101420_20050922T101442_DLR_10035_PREU.BIL.ZIP
  Length     Date   Time    Name
 --------    ----   ----    ----
222874960  04-15-08 10:55   imagery.bil
      150  04-15-08 10:55   imagery.blw
      120  04-15-08 10:55   imagery.hdr
407375360  04-15-08 10:57   geolayer.bil
     2411  04-15-08 10:57   metadata.xml
      672  04-15-08 10:54   calib.dat
  1412334  04-15-08 10:54   050922P6029032L0000S3_ssd.txt
 --------                   -------
631666007                   7 files

Directory:  2006_DLR/VER1-0/li/Coverage_2/European
Archive:
SP04_HRV1_X__1O_20060615T102622_20060615T102631_DLR_227_PREU.BIL.ZIP
  Length     Date   Time    Name
 --------    ----   ----    ----
 33876640  10-27-07 19:03   imagery.bil
      150  10-27-07 19:03   imagery.blw
      120  10-27-07 19:03   imagery.hdr
 72000000  10-27-07 19:03   geolayer.bil
     2376  01-23-08 14:00   metadata.xml
     1596  10-27-07 19:03   calib.dat
  2048190  10-27-07 19:03   metadata.dim
 --------                   -------
107929072                   7 files
```

Figure 1.2: List of files contained in an archive file: examples for IRS and SPOT.

Table 1.1: Total number of images per sensor for first and second coverages. The number between parenthesis indicates the number of unique source images. Indeed, the image for the European coverage has been received on a country basis so that some images have been orthorectified and delivered more than once. Status as of Thu Apr 24 13:53:15 CEST 2008 (39 countries).

	1st coverage	2nd coverage	combined coverages
SPOT4	861 (836)	663 (633)	1524 (1453)
SPOT5	552 (541)	340 (335)	892 (871)
IRS-P6	667 (627)	616 (593)	1283 (1209)
total	2080 (2004)	1619 (1561)	3699 (3533)

geometry and sometimes even in data footprint extent). Consequently, the data referred to as European corresponds to a union of national data made available in the ETRS-LAEA projection. There is therefore a risk that the final European mosaic will come down to a union of 38 country mosaics rather than a truly pan-European mosaic. Note that a total of 166 scenes were delivered two or three times. The former case occurred 148 times and the later 9 times (i.e., $166 = 148 + 9 \times 2$).

All duplicate scenes were processed at a later stage in such a way that the final number of 3,533 unique scenes was reached (see Sec. 1.5). This final set of unique images is referred to as the *combined* or *reference coverage*. Given the country based processing, the quality of the orthorectification of an image is only guaranteed within the country this image belongs to[1]. This feature is taken into account when determining the domain of reliability of the data masks of each image, see Sec. 1.5. Finally, table 1.2 details the number of images received per country, per sensor, and per coverage. The master list with scene names per country and coverage was made available to the JRC on the 12th of August 2008. This list is identical to the list reconstructed from the received data (apart from the 21 ni scenes that were not included in the master list).

1.2 Naming convention

The heterogeneity of the input imagery motivated us to develop a naming convention suitable for any type of sensor [40]. Therefore, it should not rely on sensor specific organisation such as the path/row numbers of Landsat and IRS satellites. A generic solution is achieved by including the x-y coordinates of the centre of the scene in the name of the file corresponding to that scene. The coordinates are given in the ETRS-LAEA projection. Files are themselves organised according to the European reference grid [2]. Figure 1.3 shows a graphical representation of the levels 18 and 19 of the European grid. Given the spatial resolution of the imagery delivered in the LAEA projection (25m), we have opted for tiles at the grid level 18. That is, each file is stored in a sub-directory whose tile name is defined by the coordinates of the centre of the corresponding scene. For example, a file whose centre coordinate belongs to the tile 42-D, is stored in the directory named `42-D`. The full file names are built by concatenating 4 underscore separated sub-fields: `YYYYMMDD-hhmm_sss_tt-iiii_xxxxxyyyyy-qq.tif`. The meaning of each sub-field is described hereafter [40]:

- `YYYYMMDD-hhmm`: reference date and time;
- `sss`: sensor acronym (e.g., SP4 for SPOT4, SP5 for SPOT5, IL3 for IRS LISS 3, and DMC for Disaster Monitoring Constellation);
- `tt-iii`: `tt` for data type composed of two letters, the first being either I (individual) or C (composite) and the second being either P (physical parameter, e.g. digital number or top of atmosphere reflectance), M (mask, e.g. segment labels or classified image), or X (mixture for P and M combined into a single multiband file) and `iii` for identifier with no restriction on length. For example, we use the identifier IP-B1XX for the

[1] A detailed assessment of the relative geometric accuracy between all pairs of overlapping scenes is presented in Chap. 4.

Table 1.2: Total number scenes received per country (country codes used in data delivered by DLR [27]), per coverage, and per sensor. Corresponding to the data sent by DLR on the 26th of February and received at JRC on the 28th of February 2008.

	coverage 1				coverage 2				combined coverages			
	SP4	SP5	IL3	ALL	SP4	SP5	IL3	ALL	SP4	SP5	IL3	ALL
al	8	1	5	**14**	2	1	4	**7**	10	2	9	**21**
at	15	5	12	**32**	11	4	15	**30**	26	9	27	**62**
ba	3	5	9	**17**	3	8	6	**17**	6	13	15	**34**
be	3	6	5	**14**	4	5	3	**12**	7	11	8	**26**
bg	24	11	12	**47**	11	8	12	**31**	35	19	24	**78**
ch	11	6	5	**22**	8	0	7	**15**	19	6	12	**37**
cs	21	11	8	**40**	9	2	12	**23**	30	13	20	**63**
cy	0	1	2	**3**	0	0	2	**2**	0	1	4	**5**
cz	10	8	8	**26**	7	3	13	**23**	17	11	21	**49**
de	48	52	44	**144**	15	7	41	**63**	63	59	85	**207**
dk	18	14	4	**36**	11	6	8	**25**	29	20	12	**61**
ee	12	3	7	**22**	0	1	8	**9**	12	4	15	**31**
es	21	23	59	**103**	15	26	65	**106**	36	49	124	**209**
fi	34	6	47	**87**	37	3	36	**76**	71	9	83	**163**
fr	95	50	47	**192**	23	10	62	**95**	118	60	109	**287**
gb	85	23	23	**131**	99	56	15	**170**	184	79	38	**301**
gr	45	30	21	**96**	12	13	34	**59**	57	43	55	**155**
hr	2	10	10	**22**	4	6	10	**20**	6	16	20	**42**
hu	20	16	5	**41**	8	1	10	**19**	28	17	15	**60**
ie	16	13	10	**39**	25	8	6	**39**	41	21	16	**78**
is	59	17	0	**76**	0	0	0	**0**	59	17	0	**76**
it	23	34	37	**94**	48	44	41	**133**	71	78	78	**227**
li	0	0	1	**1**	1	0	0	**1**	1	0	1	**2**
lt	9	17	4	**30**	11	3	7	**21**	20	20	11	**51**
lu	0	2	1	**3**	3	0	0	**3**	3	2	1	**6**
lv	18	2	6	**26**	3	1	8	**12**	21	3	14	**38**
mc	3	0	6	**9**	3	1	3	**7**	6	1	9	**16**
me	3	0	3	**6**	3	1	2	**6**	6	1	5	**12**
mt	0	0	1	**1**	0	1	0	**1**	0	1	1	**2**
ni	3	0	3	**6**	12	2	1	**15**	15	2	4	**21**
nl	11	13	5	**29**	15	8	2	**25**	26	21	7	**54**
no	62	9	57	**128**	129	32	6	**167**	191	41	63	**295**
pl	52	10	29	**91**	14	5	31	**50**	66	15	60	**141**
pt	10	5	12	**27**	19	5	10	**34**	29	10	22	**61**
ro	22	22	24	**68**	31	11	21	**63**	53	33	45	**131**
se	49	66	42	**157**	27	15	37	**79**	76	81	79	**236**
si	6	3	2	**11**	4	3	3	**10**	10	6	5	**21**
sk	12	5	5	**22**	2	3	6	**11**	14	8	11	**33**
tr	28	53	86	**167**	34	37	69	**140**	62	90	155	**307**
ALL	**861**	**552**	**667**	**2080**	**663**	**340**	**616**	**1619**	**1524**	**892**	**1283**	**3699**

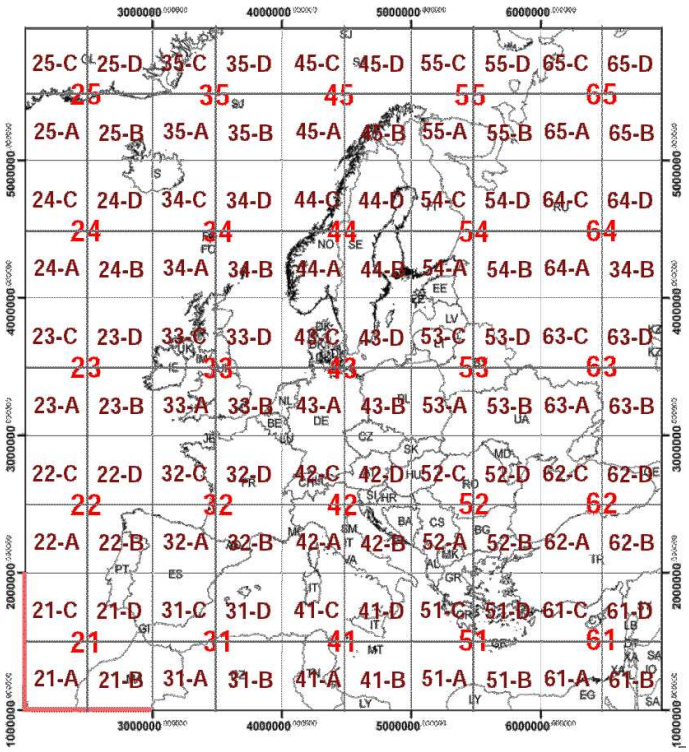


Figure 1.3: European grid level 18 and 19: grid space 500km × 500km (level 18) and 1000km × 1000km (level 19). Source: [40].

first band of a SPOT or IRS image delivered for the country with country code XX. This identifier is unique even in case the same scene was delivered for more than one country. Once identical imagery originating from more than one country or coverage are merged to create the reference coverage (see Sec. 1.6), we use the identifier IP-B10C for the first band of a SPOT or IRS image whatever the name of the country or countries it was delivered for (this information can be retrieved from the header file described in Sec. 1.4).

- `xxxxxyyyyy-qq`: reference pixel cell code indicating the position of the median pixel of the area filled with data values (see details in Sec. 1.5.1). The 5 `x` (resp. `y`) numerals are in units of 100m. They indicate the x- (resp. y-) coordinate of the lower left corner of the 100m square block in which the reference pixel falls. For example, the reference pixel cell code of a scene with centre in Sicily will always be in the form `45xxx15yyy-qq` because Sicily lies fully in the level 18 tile `41-D`, see Fig. 1.3. The two `qq` characters belong to the set $\{\texttt{A}, \texttt{B}, \texttt{C}, \texttt{D}\}$ and are used for coding the position of the reference pixel in a two-level quad-tree representation of each 100m square block. That is, the combination of the 4 letters code all 16 25m blocks. For instance, the code `AA` indicates that the reference

pixel is the lower left 25m block[2].

1.3 Storage format

For storing ingested imagery data, we have chosen the GeoTIFF [31] format given its flexibility and high portability. This format represents an effort by over 160 different remote sensing, GIS, cartographic, and surveying related companies and organisations to establish a TIFF [1] based interchange format for georeferenced raster imagery. TIFF allows for embedded lossless compression. We have opted for the LZW compression [44] with horizontal differencing. Because the GeoTIFF specifications [31] indicate that EPSG Projection System codes [5] must be in the range $[20000, 32760]$, the EPSG projection code for ETRS89-LAEA (3035) is not compliant. For this reason, all projection parameters are given in full. Figure 1.4 lists the parameters as extracted from a sample ETRS-LAEA file.

1.4 Metadata

While the GeoTIFF files contain all necessary information for successfully reading an image including all information regarding its associated coordinate reference system and position within this system, the GeoTIFF format does not yet include standardised tags for storing additional metadata such as those required for transforming the provided digital numbers into reflectance values. For this reason, it was decided to accompany each GeoTIFF file with a header file following the ENVI specifications (in ASCII format). Apart form the ENVI required fields, additional useful metadata was included. This additional information begins with a semi-colon (indicating a comment for ENVI), then the name of the field and its value. The content of a typical header file is shown in Fig. 1.5 for a SPOT image and Fig. 1.6 for an IRS image. Apart from the path and row fields only available for IRS-P6 images, these header files contain the same fields. The at-sensor solar exoatmospheric irradiance values were extracted from the delivered metadata file for the SPOT sensors[3] and from [29, 8] for the IRS-P6 sensor.

1.5 Data/no data masking

The delivered data does not contain information that could lead directly to data/no data masks for each delivered image. To address this lack of information, an ad hoc (indirect) procedure was developed. This procedure consists of two main steps. The first solves the data/no data masking per se. The second step addresses the fact that scenes were processed on a country basis (so that there is unfortunately no guarantee that a scene delivered for a country is reliable outside the corresponding country's territory).

[2]Unfortunately, the code used for generating the 2 letter code did not properly generated the second letter of the code. For instance, the letter `D` is never produced. In practice, this leads to an imprecision of 1 pixel in the actual location of the reference pixel.

[3]see also [26] and `http://www.spotimage.fr/automne_modules_files/standard/public/p229_fileLINKEDFILE_new_lumina.PDF`.

```
Geotiff_Information:
   Version: 1
   Key_Revision: 1.0
   Tagged_Information:
      ModelTiepointTag (2,3):
         0               0               0
         916475          2825275         0
      ModelPixelScaleTag (1,3):
         25              25              0
      End_Of_Tags.
   Keyed_Information:
      GTModelTypeGeoKey (Short,1): ModelTypeProjected
      GTRasterTypeGeoKey (Short,1): RasterPixelIsArea
      GTCitationGeoKey (Ascii,9): "JRC-EGCS"
      GeographicTypeGeoKey (Short,1): GCS_EUREF89
      GeogCitationGeoKey (Ascii,7): "ETRS89"
      GeogGeodeticDatumGeoKey (Short,1): Datum_European_Reference_System_1989
      GeogPrimeMeridianGeoKey (Double,1): 3.16640047e-269
      GeogLinearUnitsGeoKey (Short,1): Linear_Meter
      GeogAngularUnitsGeoKey (Short,1): Angular_Degree
      GeogEllipsoidGeoKey (Short,1): Ellipse_GRS_1980
      GeogSemiMajorAxisGeoKey (Double,1): 6378137
      GeogInvFlatteningGeoKey (Double,1): 298.257222
      GeogPrimeMeridianLongGeoKey (Double,1): 0
      ProjectedCSTypeGeoKey (Short,1): Unknown-3035
      PCSCitationGeoKey (Ascii,19): "ETRS89 / ETRS-LAEA"
      ProjectionGeoKey (Short,1): User-Defined
      ProjCoordTransGeoKey (Short,1): CT_LambertAzimEqualArea
      ProjLinearUnitsGeoKey (Short,1): Linear_Meter
      ProjFalseEastingGeoKey (Double,1): 4321000
      ProjFalseNorthingGeoKey (Double,1): 3210000
      ProjCenterLongGeoKey (Double,1): 10
      ProjCenterLatGeoKey (Double,1): 52
      End_Of_Keys.
   End_Of_Geotiff.

PCS = 3035 (ETRS89 / ETRS-LAEA)
Projection = 19986 (Europe Equal Area 2001)
Projection Method: CT_LambertAzimEqualArea
   ProjCenterLatGeoKey: 52.000000 ( 52d 0' 0.00"N)
   ProjCenterLongGeoKey: 10.000000 ( 10d 0' 0.00"E)
   ProjFalseEastingGeoKey: 4321000.000000 m
   ProjFalseNorthingGeoKey: 3210000.000000 m
GCS: 4258/ETRS89
Datum: 6258/European Terrestrial Reference System 1989
Ellipsoid: 7019/GRS 1980 (6378137.00,6356752.31)
Prime Meridian: 8901/Greenwich (0.000000/  0d 0' 0.00"E)
Projection Linear Units: 9001/metre (1.000000m)

Corner Coordinates:
Upper Left    (  916475.000, 2825275.000)  ( 31d52'24.43"W, 39d42' 3.18"N)
Lower Left    (  916475.000, 2739625.000)  ( 31d23'37.09"W, 39d 3'24.40"N)
Upper Right   ( 1000350.000, 2825275.000)  ( 31d 0' 6.36"W, 40d 6'12.09"N)
Lower Right   ( 1000350.000, 2739625.000)  ( 30d31'35.49"W, 39d27'16.09"N)
Center        (  958412.500, 2782450.000)  ( 31d11'53.78"W, 39d34'49.29"N)
```

Figure 1.4: GeoTIFF header information on a sample file 20070823-1303_SP4_IP-B10C_0958427824-CA.tif indicating all available geoinformation (output of the `listgeo` command).

```
ENVI
description = { Image2006 - ingested on 2008-05-18 based on ingestIM2K6 v1.2}
samples = 3355
lines = 3426
bands = 1
header offset = 0
file type = TIFF
data type = 1
interleave = bsq
sensor type = SPOT
byte order = 0
read procedures = {idl_tiff_read_spatial, idl_tiff_read_spectral}
map info = {ETRS-LAEA, 1.0000, 1.0000,916475.0, 2825275.0, 25.0, 25.0, ETRS89, units=Meters}
projection info = {11, 6378137.0, 6356752.3, 52.000000, 10.000000, 4321000.0, 3210000.0,
                   ETRS89, ETRS-LAEA, units=Meters}
band names = {SP4 B1GRN}
data ignore value = 0
default stretch = 2.0% linear
wavelength units = Micrometers
wavelength = {0.545}
bbl = {1}
default bands = {1}
data gain values = { 0.3344 }
data offset values = { 0 }
;sunElevation = 58.9
;sunAzimuth = 148.4
;acquisitionDate = 20070823
;acquisitionTime = 1303
;resampling = { CC }
;nominalAltitude = 830462.99673
;solarIrradianceValue = 1851
;imageNoDataFraction = 0.5
;countryOrigin = pt
;originalMetadata = 60260001_1A_DVD.ZIP
;parentFile = SP04_HRV2_X__1O_20070823T130328_20070823T130337_DLR_115_PREUpt
;dataSet = Image2006Coverage1LAEA
```

Figure 1.5: SPOT header file (20070823-1303_SP4_IP-B10C_0958427824-CA.hdr).

```
ENVI
description = { Image2006 - ingested on 2008-05-16 based on ingestIM2K6 v1.0}
samples = 7728
lines = 7746
bands = 1
header offset = 0
file type = TIFF
data type = 1
interleave = bsq
sensor type = IRS LISSIII
byte order = 0
read procedures = {idl_tiff_read_spatial, idl_tiff_read_spectral}
map info = {ETRS-LAEA, 1.0000, 1.0000,3537850.0, 2584450.0, 25.0, 25.0, ETRS89, units=Meters}
projection info = {11, 6378137.0, 6356752.3, 52.000000, 10.000000, 4321000.0, 3210000.0,
                   ETRS89, ETRS-LAEA, units=Meters}
band names = {IL3 B1GRN}
data ignore value = 0
default stretch = 2.0% linear
wavelength units = Micrometers
wavelength = {0.555}
bbl = {1}
default bands = {1}
data gain values = { 0.724353 }
data offset values = { 0 }
;maxgray = 255
;sunElevation = 54
;sunAzimuth = 156.5
;path = 19
;row = 37
;acquisitionDate = 20070419
;acquisitionTime = 1058
;nominalPixelSpacing = 23.500
;nominalLineSpacing = 23.500
;resampling = { CC }
;nominalAltitude = 817000.0000000
;solarIrradianceValue = 1849.50
;imageNoDataFraction = 0.46
;countryOrigin = fr
;originalMetadata = 070419P6019037L0000S3.ZIP
;parentFile = IR06_LI3_ORT_1D_20070419T105837_20070419T105859_DLR_18191_PREUfr
;dataSet = Image2006Coverage2LAEA
```

Figure 1.6: IRS-LISSIII header file (`20070419-1058_IL3_IP-B10C_3634924873-CB.hdr`).

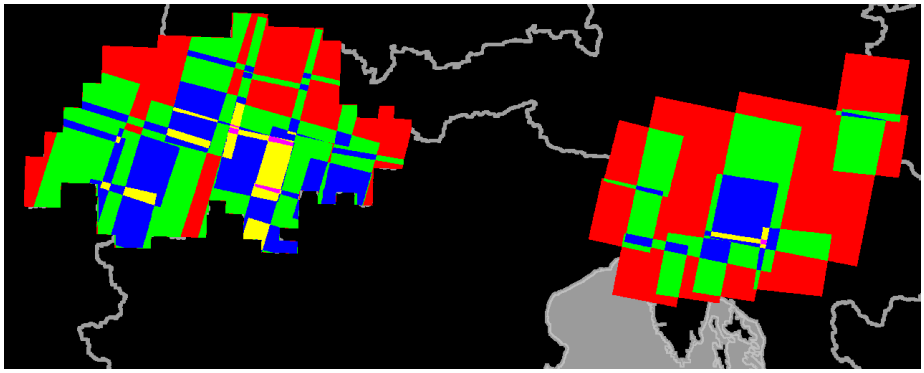


Figure 1.7: Point-wise maximum mosaic of the Swiss and Slovenian images for all images delivered for these countries in coverage 1. Colours are used for coding the degree of overlap using the order red, green, blue, yellow, magenta (i.e., no overlap in red and 5 images overlapping in magenta). Observe the non-convex data footprints in the case of Switzerland.

1.5.1 Where are data points?

Since no specific digital number (DN) was reserved for no data values and no data masks were not available, the generation of a mask indicating where actual data measurements are available can only be estimated. This is achieved with the following procedure:

- threshold each band for values ≥ 1 and sum up the resulting bands;
- threshold the resulting image for values ≥ 2. That is, a pixel is deemed to be a data point if this pixel has a DN value greater than 0 for at least two bands (theoretically, a non zero value in one band would be sufficient but this has not been tested in practice);
- fill the holes of the resulting image;
- erode the resulting image by a 3×3 square since the interpolated values along the border of this mask are corrupted by boundary effects.

The resulting data masks (also called Regions Of Interest or ROI) are named using the following data type and identifier field:

- `IM-DRXX` for delivered data where XX refers to the country code the image was delivered for;
- `IM-DROI` for the generated reference coverage.

All orthorectified images were delivered with data within their full original footprint except for Switzerland where large parts falling outside this territory were suppressed leading to non convex data footprints. The discrepancy between the Swiss data footprints and those obtained for all other countries is illustrated in Fig. 1.7 using Slovenia as example for comparison purposes.

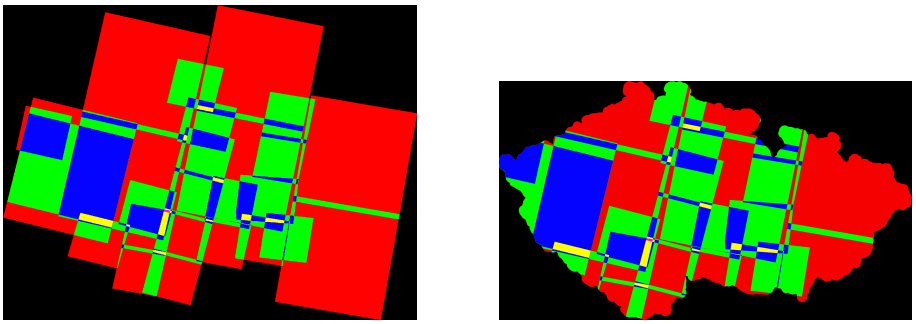


Figure 1.8: Data versus country regions of interest in the case of the Czech Republic. Left: `IM-DRCZ` Right: `IM-CRCZ`.

1.5.2 Which data points are reliable?

Reliable data can be defined as data that was processed in such a way that the errors against the target precision are within predefined limits. Due to the fact that images we processed on a country basis without ensuring that the accuracy requirements were met outside the territory of the processed country, an image originating from a given country is only geometrically reliable within the boundaries of this country.

The domain of reliability of each received image is defined by the intersection of the full footprint (`IM-DRXX` described above) with the territory of the country the image originates from. However, for generating a European mosaic we need to considered buffered territories so as to create some overlap between neighbouring countries because otherwise the European mosaic would come down to a simple union of country mosaics. The data has therefore been deemed as reliable when it falls within the country boundary enlarged by a buffer of 5km (400 pixels at a resolution of 25m). In addition, we have also included the sea and all parts falling outside the 39 participating countries (keeping in mind that the quality of the orthorectification is not secured in the non-participating countries). The resulting domains of reliability for the delivered data are referred to as `IM-CRXX`:

$$\texttt{IM-CRXX} = \texttt{IM-DRXX} \bigcap \left(\delta_B(\texttt{XX}) \bigcup \text{sea} \bigcup \overline{\cup_{\substack{\texttt{YY}=\texttt{TR} \\ \texttt{YY}=\texttt{AL}}} \texttt{YY}} \right),$$

where `XX` denotes the country code the image was delivered for, δ the dilation operator, B a disk of radius equal to 5km, and with the list of participating countries starting with the code `AL` (for Albania) and terminates with the code `TR` (for Turkey). That is, $\overline{\cup_{\substack{\texttt{YY}=\texttt{TR} \\ \texttt{YY}=\texttt{AL}}} \texttt{YY}}$ refers to the territory *not* covered by the union of the territories of the participating countries. Figure 1.8 illustrates the difference between data and country regions of interest in the case of the Czech Republic.

All ROI computations involving country boundaries were performed using National Territory Units at level 0 (Gisco version 9) rasterised at a spatial resolution of 25m in the European reference grid. The domains of reliability for the images of the reference coverage are referred to as `IM-CROI`. When they correspond to images that were delivered for only one country, they are identical to their corresponding `IM-CRXX`. A more general definition, holding for images

that were delivered for more than one country, is to define IM-CROIs as the union of the domains of reliability of all input delivered images corresponding to the considered scene:

$$\texttt{IM-CROI} = \bigcup_{\texttt{YY}} \texttt{IM-CRYY},$$

where the union spans over the set of countries for which the considered image was delivered (i.e., either 1, 2 or 3 countrie(s) for either unique, duplicate or triplicate images). An equivalent formula is used for creating the reference DROIs from the coverage 1 and 2 DRXXs. An example is detailed in the next section when describing the methodology developed for merging duplicate images when creating the reference coverage.

1.6 Combining duplicate images

The proposed methodology for combining duplicate images is first described. It is then applied to all 3,699 delivered images so as to create a reference coverage containing 3,533 unique scenes.

1.6.1 Methodology

Because images were processed and delivered on a country basis, multiple delivery of a number of scenes has occurred (up to 3 deliveries of the same scene). Differences in processing have led to differences between duplicates so that arbitrary selection of an image among the duplicates is not a viable solution. For example, Figure 1.9 shows on the top row a triplicate SPOT4 scene included in the Swedish, Finnish, and Norwegian coverages (all in coverage 2). The date and time strings of the input 3 scenes is as follows: 20060915-1051, the sensor name is SP4, and the centre pixel value is 4834251055-BC. The rightmost image of the top row shows the territory associated with each country (light grey for Sweden, black for Finland, and dark grey for Norway). The reliable data area associated with each scene are shown in the middle row (i.e., buffered country zones). They corresponding to IM-CRSE, IM-CRFI, and IM-CRNO respectively. Note that the delivered scenes were processed separately leading to discrepancies as highlighted by the zoom section (bottom row).

Therefore, original scenes that were delivered for more than one country need to be composed so as to create unique scenes (this process could be referred to as *image conflation*). A simple cut along the country boundaries is not necessarily a good solution because the cutline may be visible in the output image. A better solution is to find a cut line following image structures near the country border (near in the sense that it should lie within the buffer considered when creating the country ROIs). This is achieved by adapting the morphological compositing procedure described in [34] (the resulting methodology could be referred to as *morphological image conflation*). An example is illustrated in Fig. 1.10 and corresponds to the compositing of the triplicate imagery shown in Fig. 1.9. More precisely, regions where the country ROIs do not overlap are considered as seeds for each respective image. In our example, there are 3 seeds displayed with fundamental colours on the left of Fig. 1.10. A region growing procedure based on the watershed transformation [38, 43] is then applied to expand these

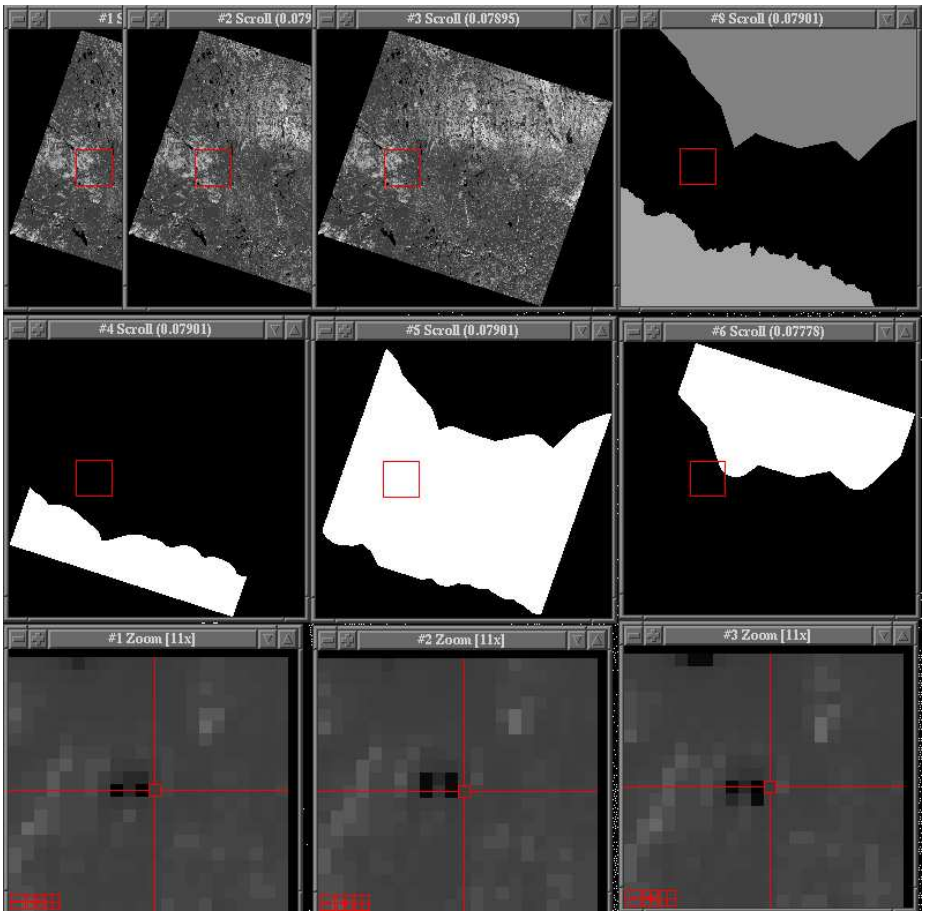


Figure 1.9: Example of a SPOT4 scene delivered for 3 countries (Sweden, Finland, and Norway) with the corresponding NUTS regions (top row), the resulting reliable data areas or `CRXX`s (national territory `XX` plus buffer of 5km away from the given boundary) (middle row), and a zoom highlighting the discrepancy between the 3 delivered images (bottom row).

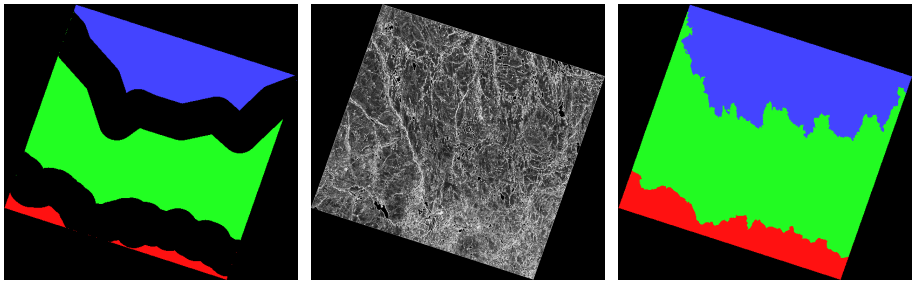


Figure 1.10: Composition of the triplicate scene shown in Fig. 1.9. Left: marker set (red for Sweden, green for Finland, and blue for Norway). Middle: point-wise minimum of the magnitude of the gradient of all three input scenes. Right: resulting decision rule obtained by propagating the markers, the propagation being driven by the gradient image.

seeds until the whole data ROI is covered, see right image of Fig. 1.10. The growth is controlled by the gradient image shown on the middle of Fig. 1.10. This gradient image is calculated by computing the point-wise minimum of the morphological gradient [32] of the 3rd band of all 3 input scenes. It ensures that resulting seam lines follow salient image structures. The grown seeds provide decision regions indicating which image should be used when composing the 3 input images. For example, the values of the Swedish images are selected in the red region. Further details on the composition procedure are given in Chap. 5. The domain of reliability of the merged scene stored in the reference coverage, i.e., `IM-CROI`, is equal to the union of the 3 country based domains of reliability shown in the middle row of Fig. 1.9. In this example, it comes down to the full data ROI of the scene.

Finally, it should be emphasised that the proposed procedure is *ad hoc* because it does not actually fix problems resulting from orthorectification performed on a country basis with different reference data and ground control points.

1.6.2 Results

Tables 1.3 and 1.4 list all duplicate and triplicate scenes with their countries and coverages of origin.

Table 1.3: The 148 duplicate scenes among all 3,699 delivered images. In case the same centre value is found, it is given only once. Similarly, if the images originate from the same cover, the coverage number is given only once.

date-time	sensor	centre(s)	country codes	coverage(s)
20050401-1043	IL3	4091728422-AA	FR_DE	2
20050404-0941	IL3	5208525144-CC	CS_RO	2
20050405-0921	IL3	5593021862-AA	GR_BG	2
20050422-1005	IL3	4738725924-DA	SI_AT	2
20050422-1005	IL3	4790528576-AA_4790428576-BB	AT_CZ	2
20050422-1006	IL3	4712224600-CC	HR_SI	2
20050502-0957	IL3	4854823406-BB	BA_HR	2
20050503-0937	IL3	5228919902-DB	GR_AL	2

date-time	sensor	centre(s)	country codes	coverage(s)
20050528-1054	IL3	3905729833-CB	BE_FR	2
20050528-1054	IL3	3952631099-BB	NL_BE	1_2
20050623-1012	IL3	4635728464-DB	AT_CZ	1_2
20050630-1024	SP5	4445144528-CB_4445144528-DA	SE_NO	1
20050701-1040	IL3	4718050208-AB_4718050207-CB	NO_SE	1
20050701-1040	IL3	4756451483-BA_4756451483-AB	NO_FI	1
20050702-1019	IL3	4936050599-DC_4936050600-AB	SE_FI	1
20050807-1206	SP4	3161436445-CC	GB_IE	1
20050830-1054	SP5	4017130899-BB	NL_BE	2
20050902-1021	SP4	4891349775-AB_4891349775-BA	SE_FI	2
20050909-1058	SP5	4459142083-DC	NO_SE	2
20051013-1147	SP5	3283635852-AB	GB_NI	2
20051020-1142	SP4	3328835513-CC	GB_NI	2
20051020-1142	SP4	3355935985-CB	NI_GB	2
20051026-0948	SP4	5065924928-CB	HR_CS	2
20051030-1150	SP4	3257136916-CB	NI_GB	2
20051030-1150	SP4	3265536307-BB	GB_NI	2
20051030-1150	SP4	3287636667-CB	GB_NI	2
20051107-1057	IL3	3533622371-BB	ES_FR	2
20051109-1014	IL3	4494225799-DC	IT_AT	2
20051109-1014	IL3	4525727112-AB	DE_AT	2
20051109-1157	SP4	3334536562-BB	GB_NI	2
20051109-1158	SP4	3304536099-AA	GB_NI	2
20060121-1153	SP4	3233935553-CB	NI_GB	2
20060221-1156	SP4	3217536303-DB	GB_NI	2
20060221-1156	SP4	3247436763-DA	NI_GB	2
20060325-0940	IL3	5149026378-AB_5149026378-AA	RO_HU	2
20060327-0925	SP4	5208219158-DC	GR_AL	2
20060506-1004	IL3	4771931167-CB	PL_CZ	2
20060510-1157	SP4	3308937149-CC	NI_GB	1
20060511-1140	IL3	3286136008-CC	NI_GB	1
20060511-1140	IL3	3354237180-DC	GB_NI	1
20060605-1123	IL3	2909320826-BC	ES_PT	2
20060605-1124	IL3	2753917659-BB	PT_ES	2
20060605-1124	IL3	2791218418-CA	ES_PT	2
20060605-1124	IL3	2850219621-CB	PT_ES	2
20060606-1102	IL3	3453422488-BB	ES_FR	1_2
20060611-1109	SP5	4433543061-AB	SE_NO	1
20060611-1109	SP5	4455843558-BB	SE_NO	1
20060612-1119	SP4	4483345985-BB_4483345985-DB	SE_NO	2
20060612-1119	SP4	4542646971-AB_4542646971-AA	SE_NO	1
20060613-1014	IL3	4499625800-DB	IT_AT	1
20060615-1026	SP4	4288627050-BA_4305927115-CC	CH_AT	1
20060620-0929	IL3	5349522733-DB	CS_BG	1
20060622-0951	SP4	4832324676-CB	BA_HR	1
20060627-0955	SP4	5107726059-CA	CS_RO	1
20060627-1134	SP4	3318436830-DC	GB_NI	1
20060701-0946	SP5	5193235272-CB	LT_PL	1
20060702-0956	SP4	5252738022-CB	LT_LV	1
20060703-0939	SP4	5194429953-DA	SK_PL	1
20060706-0952	SP5	4861024702-DB	BA_HR	1_2
20060706-1021	SP4	4684730939-DB	DE_CZ	1
20060708-0953	IL3	4926727402-CC	SK_HU	1
20060710-1052	IL3	3908829832-CB	FR_BE	1
20060711-1031	IL3	4211327084-BB_4215226885-BC	DE_CH	1_2
20060711-1032	IL3	4165025940-DC_4172525791-DB	CH_IT	1
20060712-0937	SP5	5134725559-DC_5134725560-BA	RO_CS	1
20060713-1057	SP5	4079531391-BC	DE_NL	1
20060715-1019	SP5	4148927673-AA	DE_FR	1
20060715-1046	IL3	4120033598-AA	NL_DE	1
20060715-1046	IL3	4209036146-DC	DE_DK	1_2
20060715-1047	IL3	4030031051-CA	BE_DE	1
20060715-1047	SP4	4392534980-CC	DE_DK	1_2
20060717-1005	IL3	4770431165-DB	PL_CZ	1
20060717-1006	IL3	4660925873-CB	SI_AT	1

date-time	sensor	centre(s)	country codes	coverage(s)
20060717-1011	SP4	4352827116-CA	DE_AT	1
20060717-1150	SP4	3250335691-BC	IE_GB	1
20060718-0944	IL3	5102928939-CA	SK_HU	1
20060718-0946	IL3	5005822248-DB	HR_ME	1
20060718-0950	SP4	5080329542-AC	SK_PL	1
20060719-0932	SP4	5302424194-BB	CS_BG	1_2
20060722-0944	SP5	4997525920-BB	CS_HU	1
20060722-1002	IL3	4742025924-CC	SI_AT	1
20060726-1017	IL3	4612333643-DC	DE_PL	1
20060728-0958	SP4	5139129863-CA	PL_SK	1
20060729-1050	IL3	4635651323-AA_4635651323-AB	SE_NO	1
20060803-0912	IL3	5693724759-BB	RO_BG	1
20060803-1047	IL3	4503846313-CB	NO_SE	1
20060803-1048	IL3	4459645046-CB_4459645046-CC	NO_SE	1
20060804-1031	SP5	4472942583-DA_4472942583-CB	NO_SE	1
20060806-1132	IL3	2795422440-BB	ES_PT	1
20060811-0957	SP5	5111837742-AB	LT_LV	1_2
20060811-1128	IL3	2881022233-CB	PT_ES	1
20060816-0942	IL3	5110923690-BB	CS_ME	1
20060817-0921	IL3	5528824395-CC	BG_RO	1
20060817-1014	SP4	4722428854-AA	AT_CZ	1
20060818-0956	SP4	5170422365-BB	CS_ME	1
20060819-0937	SP4	5369322130-BB	BG_MC	1_2
20060819-0937	SP4	5382223243-CB	CS_BG	1
20060819-1018	IL3	4516829710-BB	DE_CZ	1
20060820-1051	SP4	4771751853-DA_4771751853-DB	NO_FI	1_2
20060824-1008	IL3	5050752143-AC_5050652143-BB	NO_FI	1_2
20060824-1008	IL3	5073253454-CA	NO_FI	1_2
20060831-1042	SP4	4463943600-AB	NO_SE	2
20060902-0937	SP5	4934022621-CB_4934022621-DC	HR_BA	1
20060904-0946	IL3	5027523587-AB	CS_BA	1
20060904-0947	IL3	5007122251-AB	BA_HR	1_2
20060904-1040	SP5	3710521540-BC	ES_FR	1
20060905-1107	IL3	3366822610-DB	FR_ES	1_2
20060906-1000	SP5	4600825055-DB	SI_HR	1_2
20060906-1030	SP4	4047529304-AA	FR_LU	1_2
20060907-1010	SP4	4796129973-DB	PL_CZ	1
20060907-1010	SP4	4807430513-BB	PL_CZ	1
20060910-1023	SP5	4488426608-CC	IT_AT	2
20060910-1100	IL3	3809431227-DC	FR_BE	1
20060913-0958	IL3	4872428656-DB	CZ_SK	1
20060913-0959	IL3	4799424669-DB	HR_BA	1
20060914-0938	IL3	5211925154-DC	RO_CS	1
20060914-1111	SP4	4551247921-BC	NO_SE	1
20060915-1051	SP4	4689950506-CA_4689950506-CC	NO_SE	1_2
20060915-1053	SP4	4456445015-DB	SE_NO	2
20060915-1054	SP4	4446341087-AC_4446341087-AB	NO_SE	1_2
20060920-1103	SP4	3385423034-AA	ES_FR	1_2
20060922-1022	SP4	4653330917-CA	PL_CZ	1_2
20060923-1003	SP4	4851227368-AA	HU_AT	1_2
20060928-1007	SP4	5038328114-BB	SK_HU	1_2
20061004-0922	IL3	5509721695-DA	BG_GR	2
20061008-1015	SP4	4352026275-CB_4365826291-DB	CH_IT	2
20061014-1054	IL3	3624422263-AB	ES_FR	2
20061016-1011	IL3	4581025833-BA	AT_IT	2
20061016-1057	SP4	4407343541-CA_4407343541-AC	SE_NO	2
20061019-1004	SP4	5023322718-AC	BA_ME	2
20061020-0945	SP4	5331922611-AB	BG_CS	2
20061023-1206	SP4	3287435601-AC	GB_NI	2
20061103-1037	IL3	4092725816-BC_4102426004-AA	FR_CH	2
20061103-1154	SP4	3153036465-BA	GB_IE	2
20061107-1039	SP4	4175625658-BB_4186425536-AB	CH_IT	2
20070125-1158	SP4	3183935914-BA	NI_GB	2
20070203-1153	SP5	3318737125-CB	NI_GB	2
20070501-1146	IL3	3216236173-BA	NI_GB	2

Table 1.4: The 9 triplicate scenes among all 3,699 delivered images. The centre pixel identifier of the third scene is in all cases identical to that of the second scene.

date-time	sensor	centre(s)	country codes	cove-rage(s)
20050907-1028	IL3	4288127076-AA_4288127076-AA_id.	LI_DE_AT	1_1_2
20060404-1149	SP4	3225036819-CB_3225036819-CB_id.	GB_NI_IE	1_1_1
20060523-0912	IL3	5682322053-BB_5682322053-BB_id.	BG_TR_GR	1_1_2
20060615-1026	SP4	4287726548-DA_4291526585-AB_id.	CH_AT_LI	1_1_2
20060717-1145	IL3	3218236166-DB_3218236166-DB_id.	GB_IE_NI	1_1_1
20060720-1043	IL3	4060529737-AC_4060529737-AC_id.	DE_LU_FR	1_1_1
20060913-0957	IL3	4895829989-BB_4895829989-BB_id.	SK_PL_CZ	1_1_1
20060915-1051	SP4	4834251055-BC_4834251055-BC_id.	SE_FI_NO	2_2_2
20060924-0944	SP4	5154426127-DB_5154426128-BB_id.	CS_RO_HU	1_1_1

date-time	sensor	centre(s)	country codes	cove-rage(s)
20070701-1103	SP5	4150943404-CC	NO_NO	1_2
20070704-1034	SP4	5036651191-DB	FI_FI	1_2
20070806-1000	SP4	5239748109-CC	FI_FI	1_2
20070807-0940	SP4	5088551065-AB	FI_FI	1_2
20070808-1029	SP5	4898950040-BB	FI_FI	1_2
20070822-1134	SP4	4137742927-BA	NO_NO	1_2
20070929-1100	SP4	4919751637-CB	NO_FI	2
20070929-1101	SP4	4880950684-CB_4880950684-CA	FI_SE	2
20071006-1028	SP4	4441044039-AC	NO_SE	2
20071009-1110	SP4	4450242572-BA_4450242572-BC	NO_SE	2

A graphical representation of the footprints of all duplicate imagery is shown in Fig 1.11. In this image, the binary footprints of all 323 duplicate scenes have been added ($323 = 148 \times 2 + 9 \times 3$). This enables the use of colours for coding the degree of overlap: red for 1 (no overlap), green for 2 (duplicates), blue for 3 (triplicates), yellow for 4 (pair of overlapping duplicates), etc. Note that the red regions appear only for scenes delivered jointly for Switzerland and a neighbour country. This is due to the non rectangular footprints of the scenes delivered for Switzerland (see also Fig. 1.7). Also note the two SPOT footprints on the west coast of Norway corresponding to the same scenes delivered for both coverages (no intersection with a country boundary). More precisely, six scenes (4 SPOT4 and 2 SPOT5) were delivered for both coverages *and* for the very same country (all for either FI or NO). Surprisingly, two of these six duplicates are not identical. The first shows only slight intensity variations but the second shows variations in both intensity and geometry! The large white regions over Northern-Ireland are due to images delivered twice for gb and ni (19 out of 21 ni scenes were also delivered for gb). Rather than compositing, mere selection was applied to (i) the 6 scenes delivered for the same country, (ii) the 19 ni/gb pairs (selecting always the gb image), and (iii) the two cs/me pairs because the current GISCO national territory units version 9 does not distinguish Serbia (cs[4]) from Montenegro (the cs image was arbitrarily selected in both cases). The remaining 130 images that were delivered two or three times were composed using the above described composition procedure. A specific module was required for the 7 duplicates involving Switzerland due to the abnormal cuts in the data ROIs for this country (see Fig. 1.7). Its description goes beyond the scope of this report.

[4]The ISO code for the republic of Serbia is in fact RS.

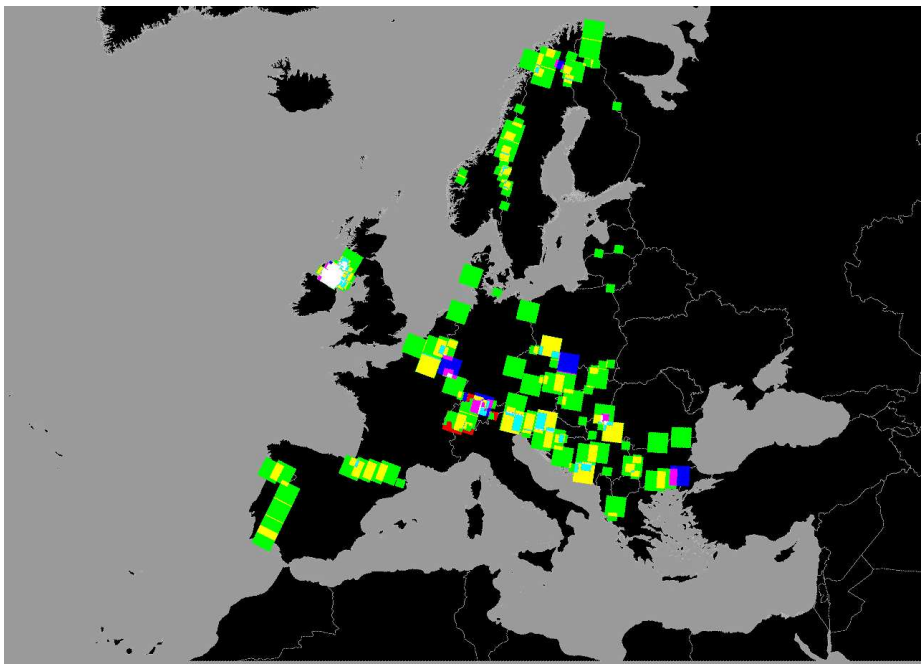


Figure 1.11: Graphical representation of the footprints of all imagery that was delivered more than once (the colours are used for coding the degree of overlap). See also tables 1.3 and 1.4.

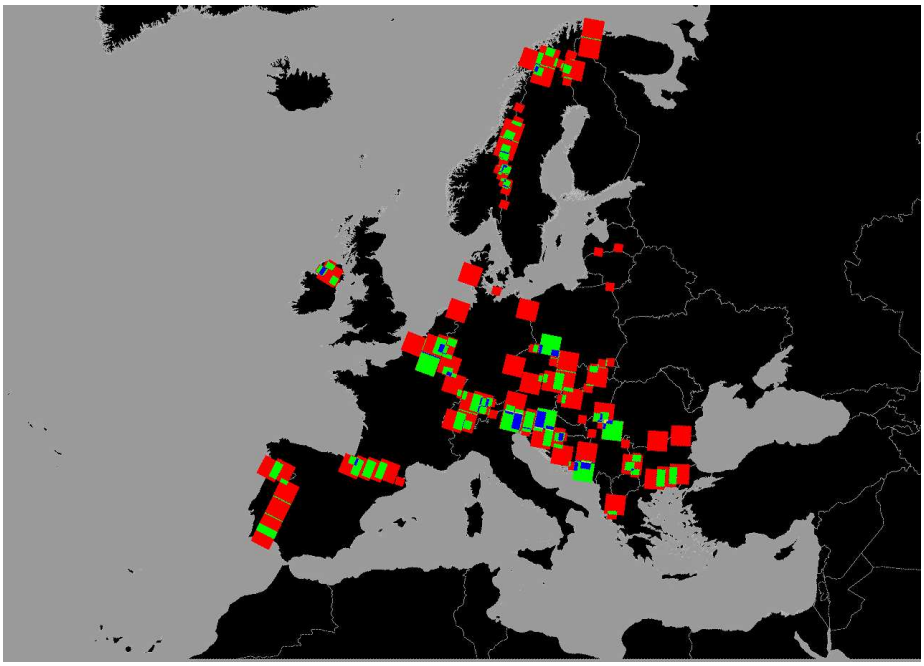


Figure 1.12: Graphical representation of the footprints of the unique scenes after composition of the duplicate imagery (the 27 scenes where mere selection occurred are not represented). Compare with Fig. 1.11.

As a result, 157 unique scenes were generated so as to avoid any duplicates in the European coverage (130 with actual composition and 27 with selection, see above). For all scenes where images were actually composed (no selection), the `countryOrigin`, `parentFile`, and `dataSet` fields of the header file contain a list of 2 (for duplicates) or 3 (for triplicates) values. A graphical representation of the footprints of all composed duplicate scenes is displayed in Fig. 1.12. The 27 scenes where mere selection occurred are not represented in this figure (i.e., the 19 ni scenes that were also delivered for gb, the 6 duplicates for the same country over the two coverages (4 for fi and 2 for no), and the 2 cs/me duplicates). For the 6 duplicates for the same country over the two coverages, the image from the first coverage has always been selected so that the `dataSet` field is always set to `Image2006Coverage1LAEA` for these images in the reference coverage. It follows that a search for all images with `dataSet` field containing `Image2006Coverage2LAEA` leads to 1,555 images instead of 1,561 as would be expected from table 1.1.

Figures 1.11 and 1.12 also reveal that only a subset of the images overlapping two or more countries were included in all these countries. Consequently, the distinction between the data and reliability domains on the basis of country extent(s) is still kept at the level of the reference coverage. The footprints of all scenes whose CROIs are not equal to their DROIs are displayed in Fig. 1.13. There are 961 such scenes (333 for SPOT4, 181 for SPOT5, and 447 for IL3). With the delivery of a truly European coverage, a one-to-one correspondence between raw images and orthorectified images would have been ensured and the reliable area would correspond to the full data footprint of the image (or at least its portion falling within the 38 participating countries).

1.7 Stored data

All ingested data described in this section is available to the IES SDI and LMNH units at the following location: `netapp2:/geodata/EUIM/`. The data is stored in 3 separate directories (`2006_COV1`, `2006_COV2`, and `2006_REF`), the first 2 for coverages 1 (2080 images) and 2 (1619 images) and the last for the merged coverages with no duplicate imagery (3533 images). Note that `2006_REF` contains symbolic links to the respective files in `2006_COV1` or `2006_COV2` except for the 130 scenes that were composed (see Sec. 1.6).

1.7.1 Coverages 1 and 2: `2006_COV1` and `2006_COV2`

The ingested received data for each coverage is stored in the directories named `2006_COV1` and `2006_COV2`. Given that the received coverages can be viewed as a union of non-harmonised country coverages rather than an actual European coverage (see Sec. 1.1), the identifier string of each generated file name contains the 2 upper case letter country code referring to the lower case country directory the data originated from. For example, `IP-B1ES` translates to band 1 originating from sub-directory named es, i.e., Spain. The full list of country codes used by DLR is as follows (see caveats in Sec. 1.1): al at ba be bg ch cs cy cz de dk ee es fi fr gb gr hr hu ie is it li lt lu lv mc me mt ni nl no pl pt ro se si sk tr.

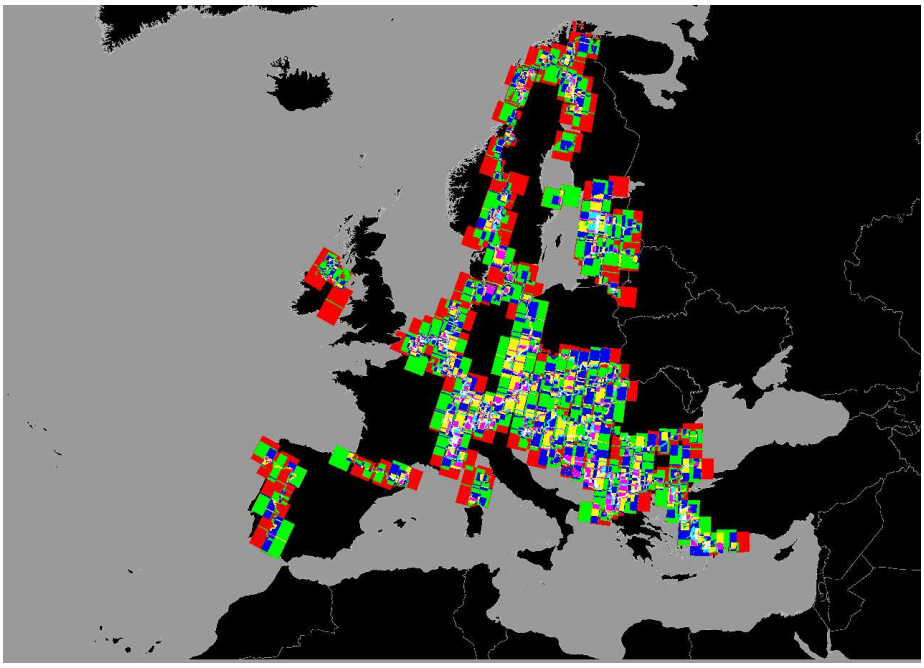


Figure 1.13: Graphical representation of all 961 CROIs of the reference coverage (total of 3,533 CROIs) that are not equal to their corresponding DROIs.

1.7.2 Combined coverages: `2006_REF`

The two input coverages are combined so as to produce a reference data set in the directory named `2006_REF`. The identifier string in the reference data set does not contain anymore the 2 letter country code used for coverages 1 and 2. For example, `IP-B1ES` becomes `IP-B10C` (first band, the letter C indicating that cubic convolution was used when interpolating the raw image during orthorectification). Data duplication is avoided for all those 3,376 images that were delivered only once by creating symbolic links to the files already stored in `2006_COV1` or `2006_COV2`. Consequently, there are only 130 images physically stored in the directory named `2006_REF` since mere selection was considered for 27 (out of a total of 157) duplicate scenes so that links were also produced for them.

Chapter 2

Coverage representation and extent

Pierre Soille, Conrad Bielski, and Joanna Nowak

Spatial Data Infrastructure Unit
Institute for Environment and Sustainability
DG Joint Research Centre, European Commission

2.1 Introduction

This chapter provides detailed representations of the footprints of the imagery delivered for the first, second, and combined coverages. The extent and degree of overlap are also measured since their are essential for the production of the mosaic. All measurements are given for both data and country based ROIs resampled by pixel selection to a resolution of 250m and using the National Territory Units at level 0 (GISCO version 9) rasterised at 250m, see Fig. 2.1.

2.2 First coverage

Figures 2.2–2.4 illustrate the regions covered by the SPOT4, SPOT5, and IRS satellites respectively. The colour are used for coding the number of images overlapping a given area. The increasing order of overlap is mapped into the fundamental colour order and their combinations (white being used for more than 7 or more overlapping image). From Fig. 2.4, it can be seen that most of target territory is covered by the IRS satellite. Noticeable exceptions include Iceland and all Atlantic Islands (beyond a series of gaps the major ones being in Scotland, Brittany, South Sweden, and the South-East corner of Turkey). The use of SPOT4 and SPOT5 imagery is mainly for covering these gaps (and cloudy regions covered by IRS), see Figs. 2.2 and 2.3. Figure 2.5 shows the degree of overlap when taking into account all three sensors. The maximum number of

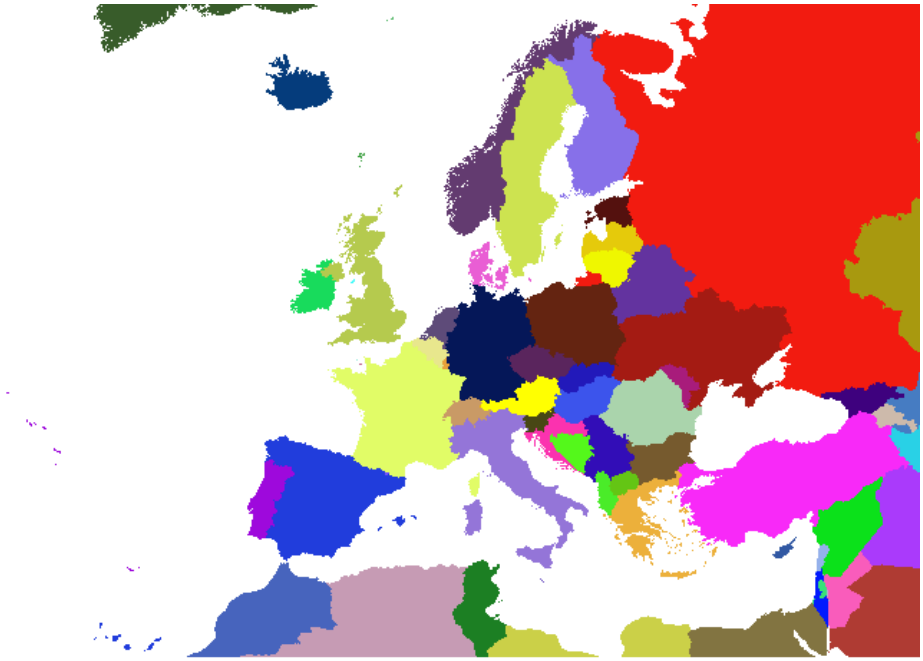


Figure 2.1: GISCO version 9 NUTS0 vector file rasterised at 250m using ETRS-89 LAEA projection and the European grid. Serbia (code RS) and Montenegro (code ME) still appear within the same unit (dark blue region in the Balkan Peninsula).

times a pixel is covered is equal to 10. Figure 2.6 shows a colour composition of the coverage by all 3 satellites.

Statistics showing the percentage of IMAGE-2006 participating country[1] covered by each sensor and all sensors together are given in tables 2.1–2.2, the first for full data ROIs (i.e., DROIs) and the second for their reduction to the territory of the country the input image originates from (i.e., CROIs). Although statistics per sensor varies when usign CROIs rather than DROIs, only in one case a larger gap occurs when considering CROIs. This happens for the Greek Megisti Island not covered by the images delivered for Greece but covered by imagery delivered for Turkey. However, if cloud free subsets of the ROIs would be considered, higher coverage percentages would be observed for other countries when using DROIs rather than CROIs, see details in Chap. 3.

Without taking cloud cover into account, tables 2.1–2.2 show that most of the target territory is covered (total gap area based on DROIs accounting for 0.0015% of the whole target territory). The largest gaps corresponds to the Porto Santo Island (42.8km^2), Madeira, Portugal. Smaller gaps can be found in Belgium (100 pixels gap), Denmark (Islet (Christianso), off Bornholm), Finland (20700 pixels in North of Finland), France (Belgium border gap and Islet in Glénan Islands), Iceland (a few coastal pixels), Italy (Islet in front of Ventotene Island), Sweden (Islet in the North Baltic sea), and UK (Saint Kilad Island).

[1]Coverage measurements for all countries at leas partly covered by IMAGE-2006 imagery are given in appendix C.

Note that the analysis is based on NUTS0 codes. Figures 2.7 and 2.9 indicate the positions of all these gaps in case of DROIs.

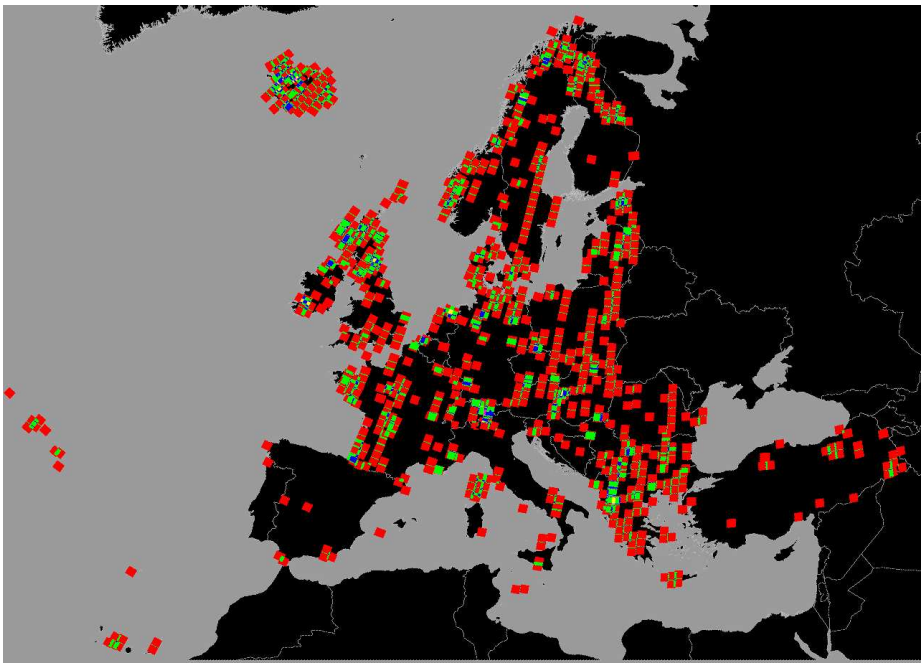


Figure 2.2: SPOT 4 footprints (first coverage) based on data ROIs. The colours are used for coding the number n of images covering a given pixel: red ($n = 1$), green ($n = 2$), blue ($n = 3$), yellow ($n = 4$), magenta ($n = 5$), cyan ($n = 6$). $n_{\max} = 6$.

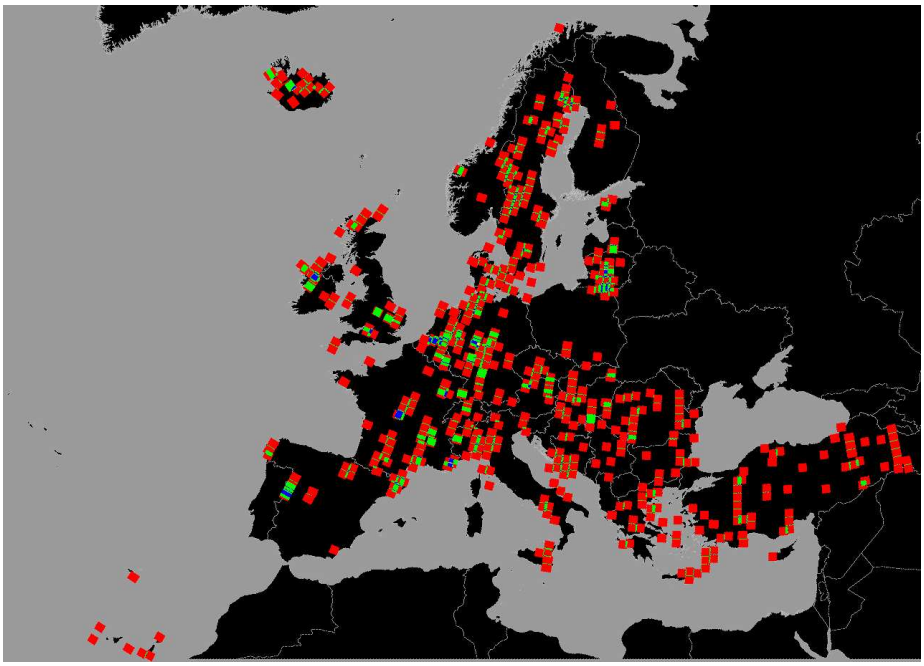


Figure 2.3: SPOT 5 footprints (first coverage) based on data ROIs. The colours are used for coding the number n of images covering a given pixel: red ($n = 1$), green ($n = 2$), blue ($n = 3$), yellow ($n = 4$). $n_{\max} = 4$. Land (resp. sea) regions not covered are painted in black (resp. light grey)

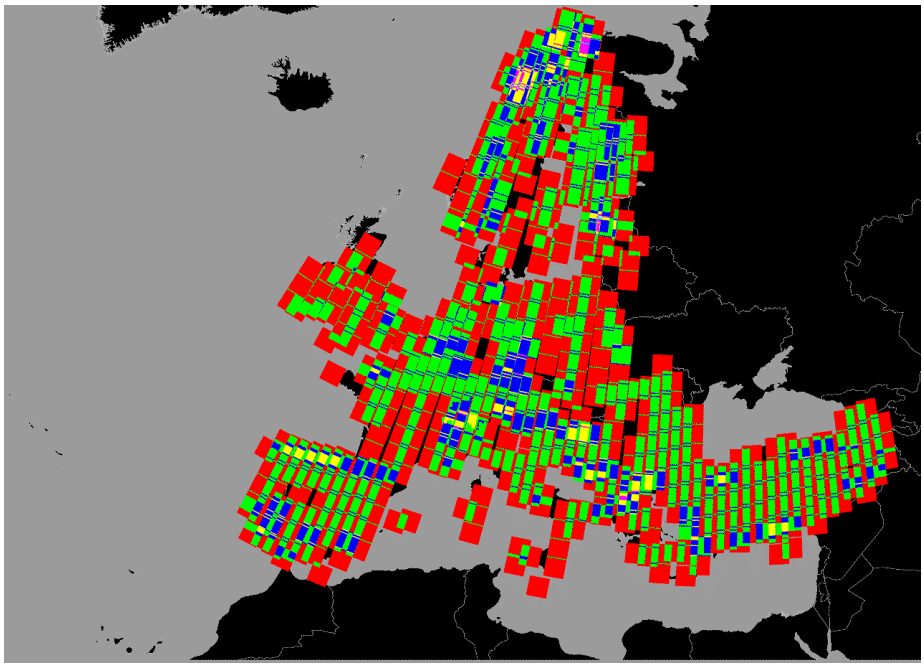


Figure 2.4: IRS-P6 LISS III-3 footprints (first coverage) based on data ROIs. The colours are used for coding the number n of images covering a given pixel: red ($n = 1$), green ($n = 2$), blue ($n = 3$), yellow ($n = 4$), magenta ($n = 5$), cyan ($n = 6$), and white $n \geq 7$. $n_{\max} = 7$. Land (resp. sea) regions not covered are painted in black (resp. light grey).

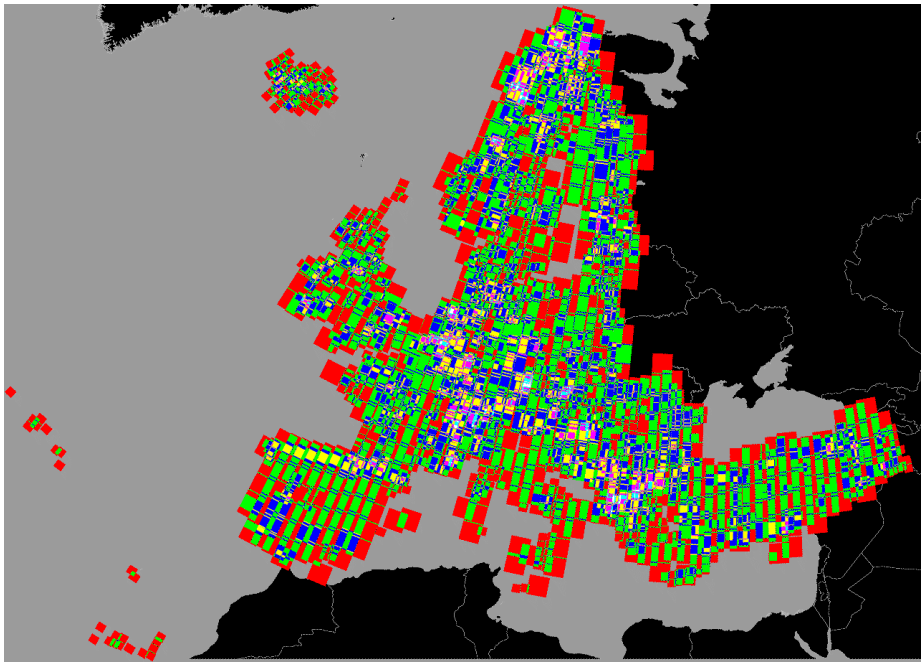


Figure 2.5: IMAGE 2006 first coverage (all sensors together). The colours are used for coding the number n of images covering a given pixel: red ($n = 1$), green ($n = 2$), blue ($n = 3$), yellow ($n = 4$), magenta ($n = 5$), cyan ($n = 6$), and white ($n \geq 7$). $n_{\max} = 9$.

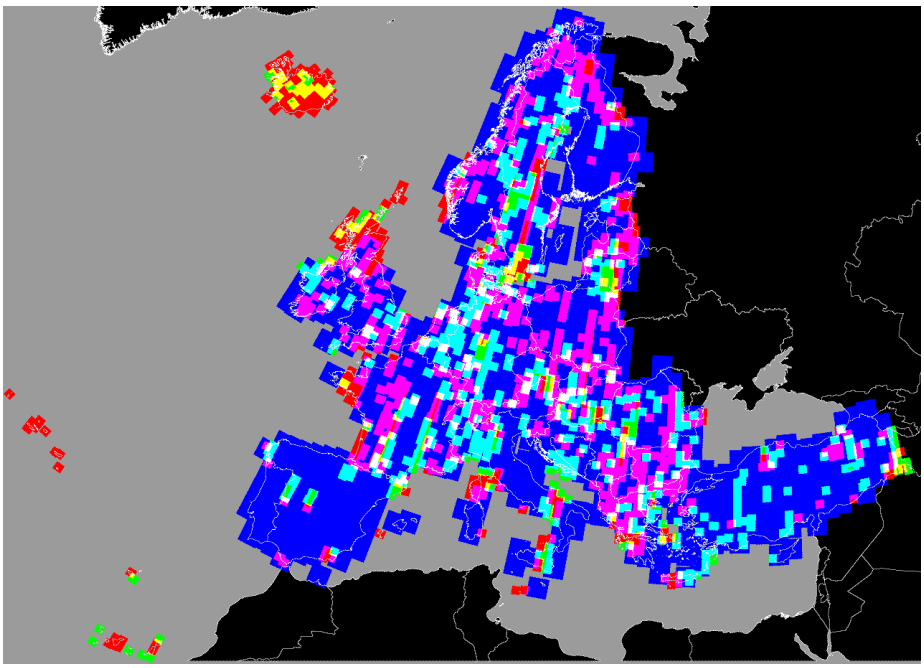


Figure 2.6: IMAGE 2006 first coverage. The colours are used for coding the type of satellite covering a given area: red for SPOT 4, green for SPOT 5, blue for IRS. The combinations of these 3 fundamental colours indicate the type of combination, e.g., yellow for areas covered both by SPOT 4 and SPOT 5 and white for regions covered by all 3 satellites. For example, regions not covered by IRS appear in either red, green or yellow. Land (resp. sea) regions not covered are painted in black (resp. light grey).

Table 2.1: First coverage statistics based on data regions of interest. The code CS refers to the former union of Serbia (RS) and Montenegro (ME).

Code[a]	area[b] (km^2)	% with SP4	% with SP5	% with IL3	% with ALL	# missing pixels	Full coverage?
AL	28534	74.77	12.18	99.82	100.00	0	yes
AT	83930	51.47	27.66	92.59	100.00	0	yes
BA	51274	18.62	34.74	100.00	100.00	0	yes
BE	30664	28.15	63.79	100.00	≈100.00	100	no
BG	110796	61.25	27.10	100.00	100.00	0	yes
CH	41288	60.49	40.74	98.91	100.00	0	yes
CS	102572	45.29	31.28	93.25	100.00	0	yes
CY	9247	0.00	8.39	100.00	100.00	0	yes
CZ	78866	41.41	30.10	92.97	100.00	0	yes
DE	357672	34.76	39.99	96.88	100.00	0	yes
DK	43360	63.33	55.94	71.45	≈100.00	200	no
EE	45327	44.70	14.81	86.22	100.00	0	yes
ES	505984	6.96	12.01	95.28	100.00	0	yes
FI	337788	30.09	6.69	97.07	≈100.00	20700	no
FR	549164	39.99	23.65	92.24	≈100.00	3700	no
GR	132025	59.90	34.51	87.91	≈100.00	1300	no
HR	56631	18.22	32.96	99.62	≈100.00	500	no
HU	93016	49.89	50.51	86.91	100.00	0	yes
IE	70175	49.98	32.24	98.39	100.00	0	yes
IS	102906	94.31	45.07	0.00	≈100.00	200	no
IT	301429	15.52	29.21	93.00	≈100.00	500	no
LI	160	100.00	0.00	100.00	100.00	0	yes
LT	64890	49.75	52.75	73.59	≈100.00	3600	no
LU	2595	37.15	77.78	100.00	100.00	0	yes
LV	64603	64.40	14.08	89.47	100.00	0	yes
MK	25153	55.58	0.88	100.00	100.00	0	yes
MT	314	0.00	0.00	100.00	100.00	0	yes
NL	37357	48.61	59.24	98.98	100.00	0	yes
NO	323456	41.05	6.59	98.47	≈100.00	100	no
PL	311894	42.47	8.68	98.14	100.00	0	yes
PT	92140	5.33	3.42	96.23	99.95	68500	no
RO	237938	32.49	32.87	93.42	100.00	0	yes
SE	449446	29.95	40.04	87.02	≈100.00	2200	no
SI	20272	56.64	12.24	100.00	100.00	0	yes
SK	49024	64.86	34.42	96.31	100.00	0	yes
TR	780122	11.61	21.66	98.18	100.00	0	yes
UK	244706	55.11	17.83	83.86	99.99	26000	no
ALL	5836738	33.51	24.54	92.34	≈100.00	127500.0	no

[a] Normalised country codes, all 38 participating countries are typeset in bold.
[b] Area of country calculated from GISCO version 9 NUTS vector file rasterised at 250m using ETRS-89 LAEA projection and the European grid.

Table 2.2: First coverage statistics based on country regions of interest. The code CS refers to the former union of Serbia (RS) and Montenegro (ME).

Code[a]	area[b] (km^2)	% with SP4	% with SP5	% with IL3	% with ALL	# missing pixels	Full coverage?
AL	28534	66.39	7.29	99.82	100.00	0	yes
AT	83930	40.81	16.68	88.27	100.00	0	yes
BA	51274	18.18	28.16	100.00	100.00	0	yes
BE	30664	17.08	51.70	100.00	≈100.00	100	no
BG	110796	57.25	26.54	92.02	100.00	0	yes
CH	41288	52.88	36.76	83.22	100.00	0	yes
CS	102572	45.05	30.14	87.17	100.00	0	yes
CY	9247	0.00	8.39	100.00	100.00	0	yes
CZ	78866	34.32	29.25	86.80	100.00	0	yes
DE	357672	30.86	37.72	96.82	100.00	0	yes
DK	43360	61.52	55.23	53.27	≈100.00	200	no
EE	45327	44.67	14.81	86.21	100.00	0	yes
ES	505984	6.52	10.66	94.17	100.00	0	yes
FI	337788	27.65	6.10	97.07	≈100.00	20700	no
FR	549164	39.43	23.54	90.85	≈100.00	3700	no
GR	132021	59.03	31.49	82.20	≈100.00	7100	no
HR	56631	8.79	23.80	93.21	≈100.00	500	no
HU	93016	43.42	47.01	73.73	100.00	0	yes
IE	70175	45.72	32.24	98.39	100.00	0	yes
IS	102906	94.31	45.07	0.00	≈100.00	200	no
IT	301429	14.01	27.57	92.99	≈100.00	500	no
LI	160	100.00	0.00	100.00	100.00	0	yes
LT	64890	37.05	52.66	72.46	≈100.00	3600	no
LU	2595	9.71	77.74	100.00	100.00	0	yes
LV	64603	62.72	10.11	85.61	100.00	0	yes
MK	25153	21.39	0.24	100.00	100.00	0	yes
MT	314	0.00	0.00	100.00	100.00	0	yes
NL	37357	47.19	58.39	92.92	100.00	0	yes
NO	323456	40.05	5.79	98.47	≈100.00	100	no
PL	311894	41.34	8.37	97.76	100.00	0	yes
PT	92140	3.97	3.42	95.70	99.95	68500	no
RO	237938	29.17	30.27	93.42	100.00	0	yes
SE	449446	28.30	39.68	82.24	≈100.00	2200	no
SI	20272	49.67	11.78	94.87	100.00	0	yes
SK	49024	49.51	31.34	83.28	100.00	0	yes
TR	780122	10.88	21.36	98.18	100.00	0	yes
UK	244706	55.11	17.83	82.92	99.99	26000	no
ALL	5836734	31.22	23.25	90.40	≈100.00	133100.0	no

[a] Normalised country codes, all 38 participating countries are typeset in bold.
[b] Area of country calculated from GISCO version 9 NUTS vector file rasterised at 250m using ETRS-89 LAEA projection and the European grid.

Figure 2.7: Position of the gaps in coverage 1 (part 1/3): BE, FR, DK, and FI.

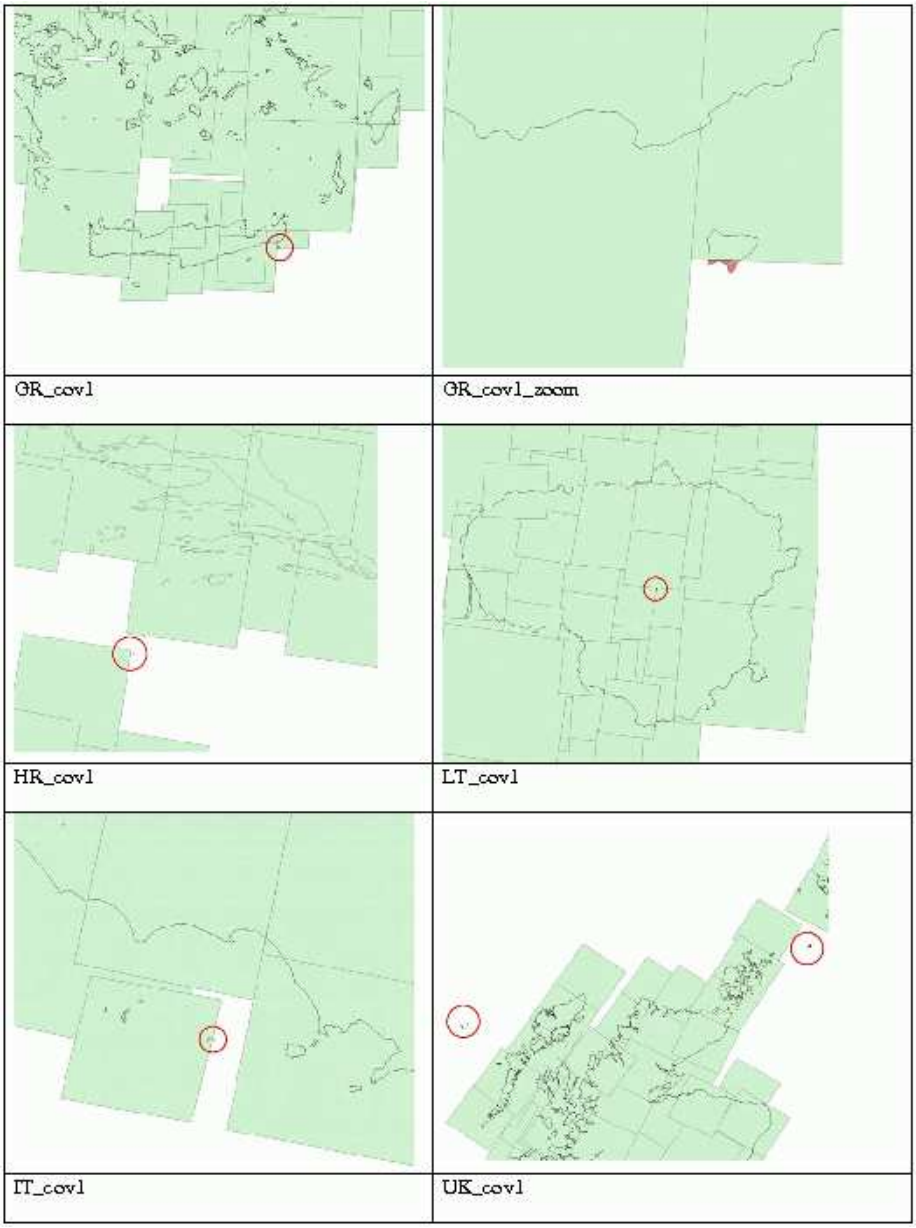


Figure 2.8: Position of the gaps in coverage 1 (part 2/3): GR, HR, LT, IT, and UK.

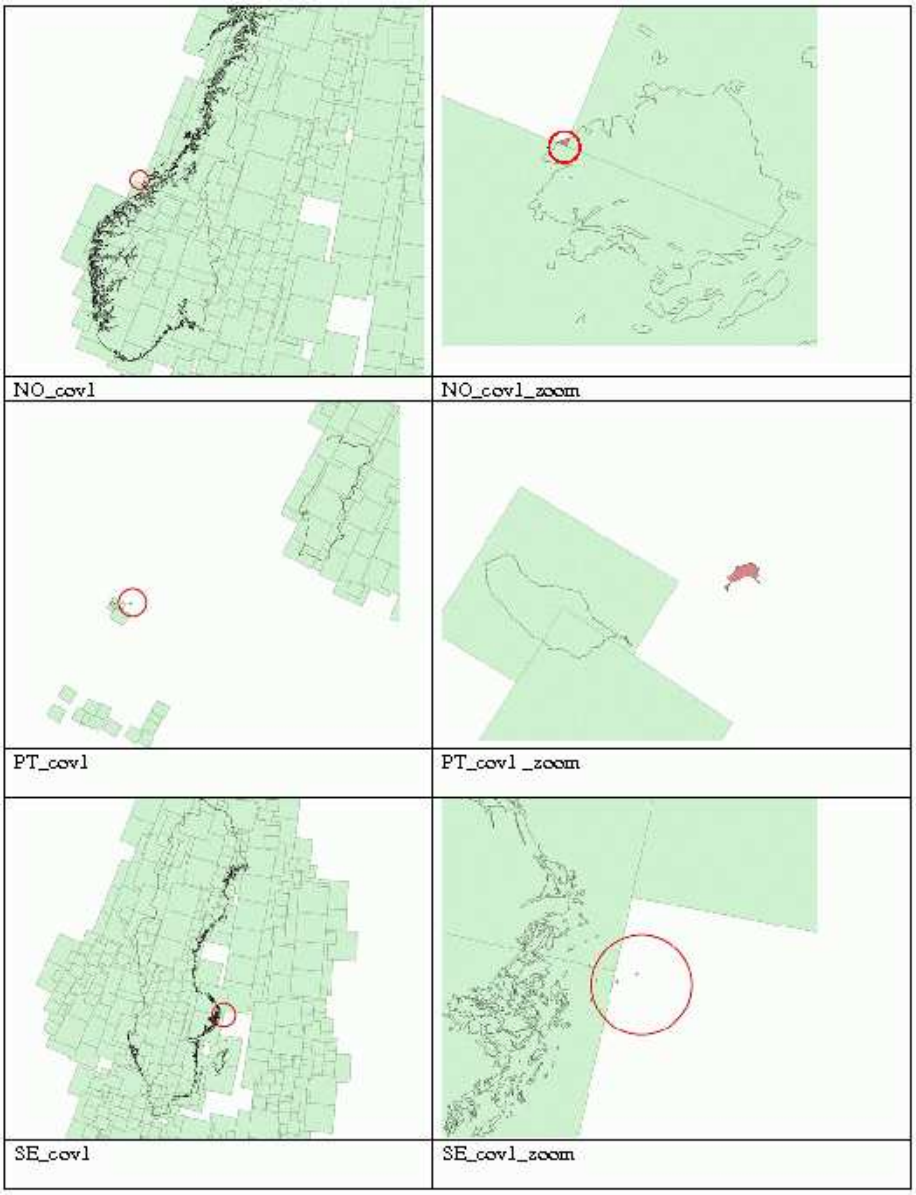


Figure 2.9: Position of the gaps in coverage 1 (part 3/3): NO, PT, and SE.

2.3 Second coverage

Figures 2.10–2.12 illustrate the regions covered by the SPOT4, SPOT5, and IRS satellites respectively. The colour are used for coding the number of images overlapping a given area. The increasing order of overlap is mapped into the fundamental colour order and their combinations (white being used for more than 6 overlapping image). In accordance with the first coverage, it can be seen that most of target territory is covered by this sensor, see Fig. 2.12. Figure 2.13 shows the degree of overlap when taking into account all three sensors. The maximum number of times a pixel is covered is equal to 12. Figure 2.14 shows a colour composition of the coverage by all 3 satellites. This image further illustrates that most of the target territory is covered by IRS.

Statistics for showing the percentage of country covered by each sensor and all sensors together are given in tables 2.3 and 2.4, the first for full data ROIs (i.e., DROIs) and the second for their reduction to the territory of the country the input image originates from (i.e., CROIs). Note that Iceland is not covered at all by the scond coverage. There are also several large gaps in Finland, Norway, and Sweden. Note that all these gaps are substantially largers when considering CROIs instead of DROIs.

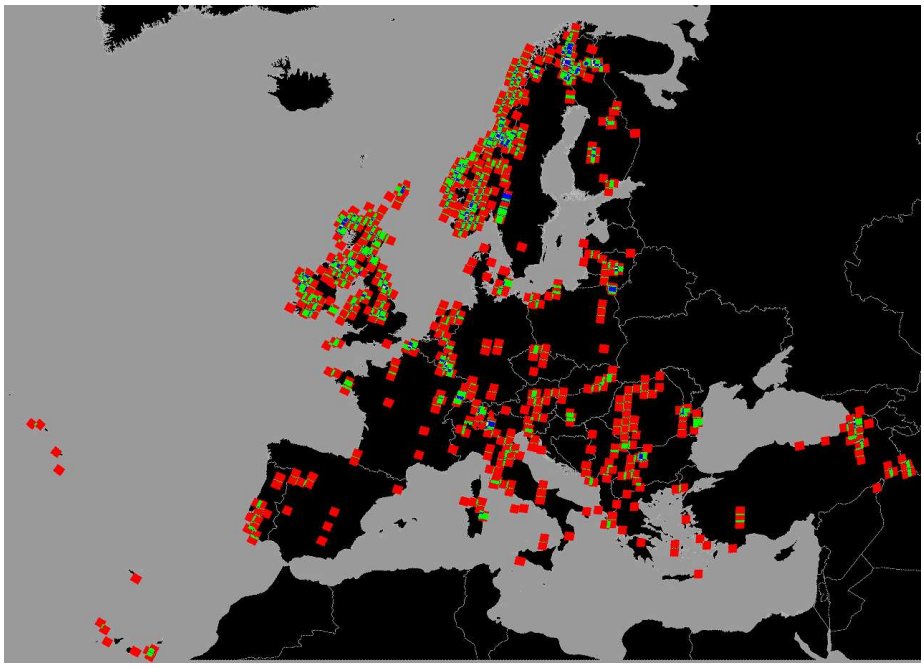


Figure 2.10: SPOT 4 footprints (second coverage) based on data ROIs. The colours are used for coding the number n of images covering a given pixel: red ($n = 1$), green ($n = 2$), blue ($n = 3$), yellow ($n = 4$), magenta ($n = 5$). $n_{\max} = 5$.

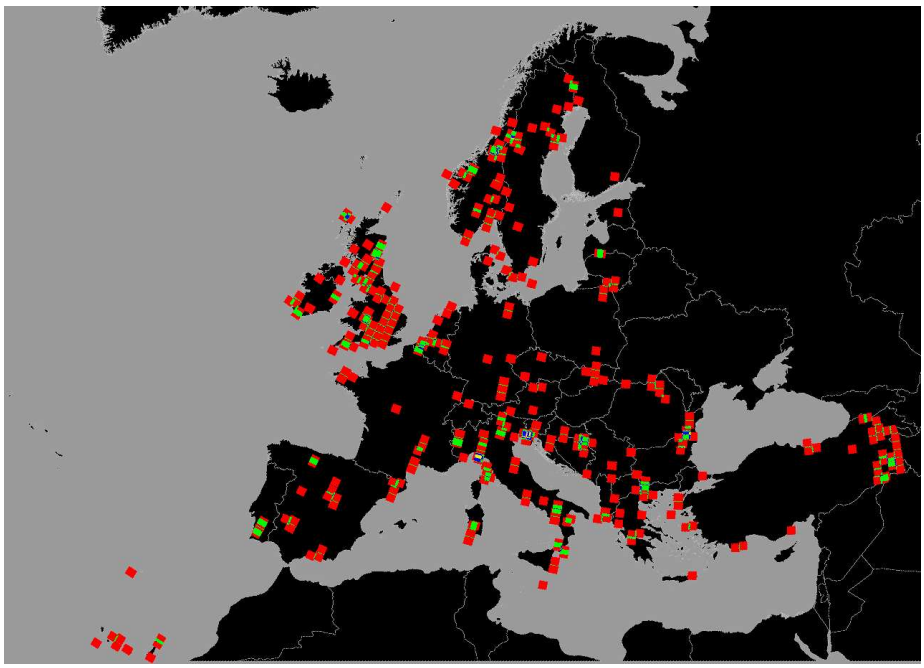


Figure 2.11: SPOT 5 footprints (second coverage) based on data ROIs. The colours are used for coding the number n of images covering a given pixel: red ($n = 1$), green ($n = 2$), blue ($n = 3$), yellow ($n = 4$), magenta ($n = 5$). $n_{\max} = 5$. Land (resp. sea) regions not covered are painted in black (resp. light grey)

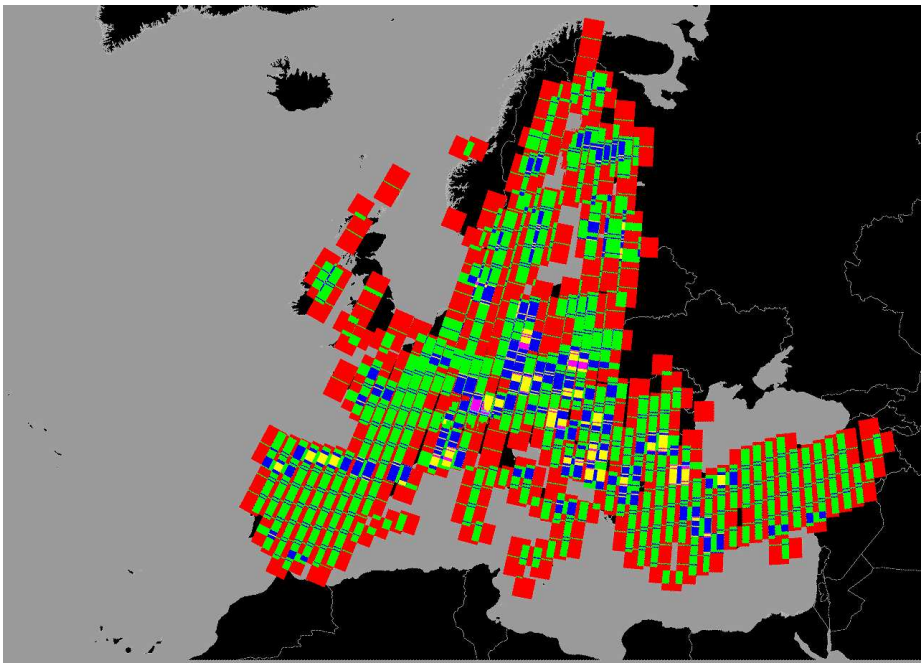


Figure 2.12: IRS-P6 LISS III-3 footprints (second coverage) based on data ROIs. The colours are used for coding the number n of images covering a given pixel: red ($n = 1$), green ($n = 2$), blue ($n = 3$), yellow ($n = 4$), magenta ($n = 5$), cyan ($n = 6$), and white $n \geq 7$. $n_{\max} = 7$. Land (resp. sea) regions not covered are painted in black (resp. light grey).

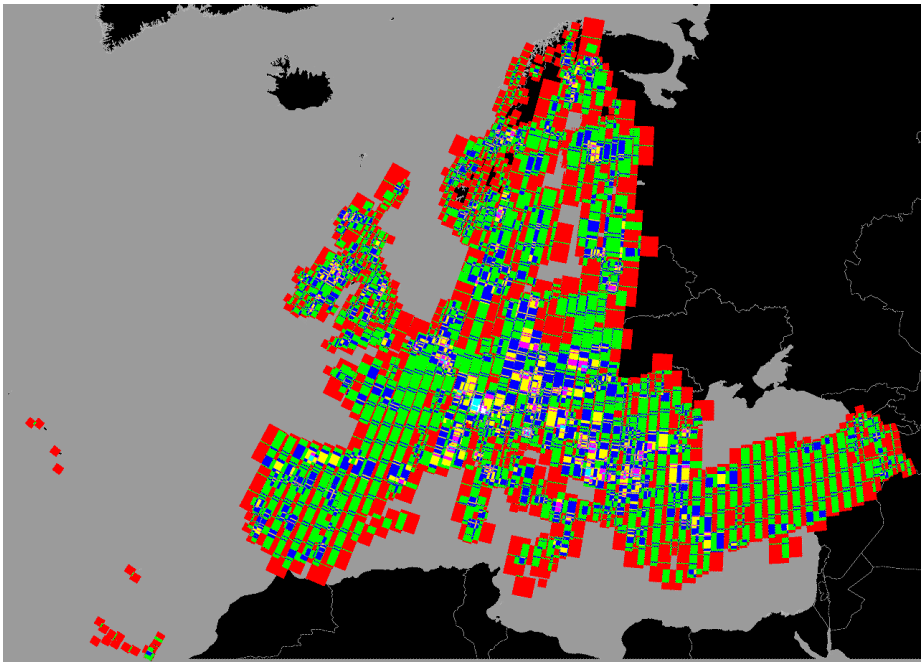


Figure 2.13: IMAGE 2006 second coverage (all sensors together). The colours are used for coding the number n of images covering a given pixel: red ($n = 1$), green ($n = 2$), blue ($n = 3$), yellow ($n = 4$), magenta ($n = 5$), cyan ($n = 6$), and white ($n \geq 7$). $n_{\max} = 9$.

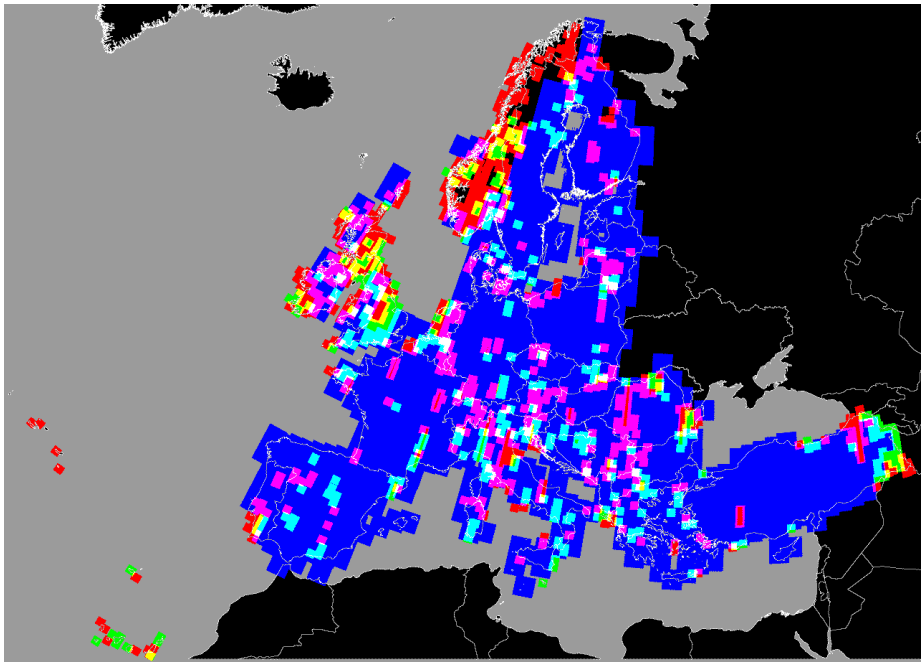


Figure 2.14: IMAGE 2006 second coverage. The colours are used for coding the type of satellite covering a given area: red for SPOT 4, green for SPOT 5, blue for IRS. The combinations of these 3 fundamental colours indicate the type of combination, e.g., yellow for areas covered both by SPOT 4 and SPOT 5 and white for regions covered by all 3 satellites. For example, regions not covered by IRS appear in either red, green or yellow. Land (resp. sea) regions not covered are painted in black (resp. light grey).

Table 2.3: Second coverage statistics based on data regions of interest. The code CS refers to the former union of Serbia (RS) and Montenegro (ME). Note that Iceland (IS) is not appearing in this table since no imagery is available in the second coverage for this country.

Code[a]	area[b] (km^2)	% with SP4	% with SP5	% with IL3	% with ALL	# missing pixels	Full coverage?
AL	28534	26.83	23.91	96.93	100.00	0	yes
AT	83930	44.10	15.46	100.00	100.00	0	yes
BA	51274	12.95	34.82	98.82	100.00	0	yes
BE	30664	34.68	39.37	99.12	100.00	0	yes
BG	110796	17.42	20.19	97.24	100.00	0	yes
CH	41288	45.90	6.85	99.99	100.00	0	yes
CS	102572	40.95	12.13	98.50	100.00	0	yes
CY	9247	0.00	0.00	100.00	100.00	0	yes
CZ	78866	22.69	12.76	99.83	100.00	0	yes
DE	357672	15.25	7.86	99.65	100.00	0	yes
DK	43360	38.20	13.77	99.43	100.00	0	yes
EE	45327	0.00	8.00	100.00	100.00	0	yes
ES	505984	7.66	11.98	98.07	100.00	0	yes
FI	329895	23.49	3.66	90.49	97.66	12649900	no
FR	549167	11.77	6.25	97.81	100.00	0	yes
GR	132025	15.65	22.83	97.04	≈100.00	100	no
HR	56631	10.58	29.71	98.09	100.00	0	yes
HU	93016	26.95	3.70	97.06	100.00	0	yes
IE	70175	70.51	24.09	72.15	100.00	0	yes
IT	301429	31.11	27.80	92.84	100.00	0	yes
LI	160	100.00	0.00	100.00	100.00	0	yes
LT	64892	44.22	16.49	100.00	100.00	0	yes
LU	2595	100.00	0.00	97.55	100.00	0	yes
LV	64603	13.99	4.27	99.96	100.00	0	yes
MK	25153	38.80	17.50	99.92	100.00	0	yes
MT	314	0.00	100.00	0.00	100.00	0	yes
NL	37357	69.17	39.92	72.21	100.00	0	yes
NO	274961	71.48	22.15	20.59	85.01	77592300	no
PL	311894	12.46	4.36	98.11	100.00	0	yes
PT	91223	36.15	11.80	89.30	98.96	1535000	no
RO	237938	37.98	12.73	87.44	100.00	0	yes
SE	402998	15.87	13.36	78.71	89.67	74319400	no
SI	20272	44.60	33.99	96.61	100.00	0	yes
SK	49024	18.50	24.78	98.99	100.00	0	yes
TR	780122	9.20	9.56	93.68	100.00	0	yes
UK	244699	56.37	50.08	59.50	99.99	36100	no
ALL	5630077	23.61	14.21	86.77	96.46	3.30784e+08	no

[a] Normalised country codes, all 38 participating countries are typeset in bold.
[b] Area of country calculated from GISCO version 9 NUTS vector file rasterised at 250m using ETRS-89 LAEA projection and the European grid.

Table 2.4: Second coverage statistics based on country regions of interest. The code CS refers to the former union of Serbia (RS) and Montenegro (ME). Note that Iceland (IS) is not appearing in this table since no imagery is available in the second coverage for this country.

Code[a]	area[b] (km^2)	% with SP4	% with SP5	% with IL3	% with ALL	# missing pixels	Full coverage?
AL	28534	10.89	12.82	96.93	100.00	0	yes
AT	83930	37.44	15.42	100.00	100.00	0	yes
BA	51274	12.95	29.79	88.14	100.00	0	yes
BE	30664	29.73	26.99	80.86	100.00	0	yes
BG	110796	14.14	19.41	94.52	100.00	0	yes
CH	41288	41.81	1.64	96.91	100.00	0	yes
CS	102572	32.54	9.55	97.64	100.00	0	yes
CY	9247	0.00	0.00	100.00	100.00	0	yes
CZ	78866	22.69	12.14	98.25	100.00	0	yes
DE	357672	12.65	6.99	99.63	100.00	0	yes
DK	43360	38.20	13.77	94.07	100.00	0	yes
EE	45327	0.00	8.00	100.00	100.00	0	yes
ES	505984	6.58	12.07	98.07	100.00	0	yes
FI	329717	22.65	2.64	90.43	97.61	12933500	no
FR	549165	10.52	5.88	97.79	≈100.00	1900	no
GR	132025	13.29	22.10	97.04	≈100.00	100	no
HR	56631	8.14	16.19	95.97	100.00	0	yes
HU	93016	22.58	2.99	94.85	100.00	0	yes
IE	70175	63.34	19.25	72.15	100.00	0	yes
IT	301429	30.73	27.74	92.59	100.00	0	yes
LI	160	100.00	0.00	100.00	100.00	0	yes
LT	64892	41.83	12.16	97.94	100.00	0	yes
LU	2595	100.00	0.00	35.81	100.00	0	yes
LV	64603	9.30	3.69	99.89	100.00	0	yes
MK	25153	19.09	8.34	99.76	100.00	0	yes
MT	314	0.00	100.00	0.00	100.00	0	yes
NL	37357	69.17	33.78	45.01	100.00	0	yes
NO	260664	69.63	22.15	17.27	80.59	100468000	no
PL	311894	11.37	3.69	97.15	100.00	0	yes
PT	91223	36.15	11.80	88.82	98.96	1535000	no
RO	237938	36.85	12.43	85.11	100.00	0	yes
SE	393877	13.93	10.07	78.47	87.64	88913800	no
SI	20272	41.51	32.86	75.30	100.00	0	yes
SK	49024	12.08	16.68	98.91	100.00	0	yes
TR	780122	8.97	9.56	93.50	100.00	0	yes
UK	244699	56.37	50.08	58.01	99.99	36100	no
ALL	5606480	21.88	13.07	85.61	96.05	3.6854e+08	no

[a] Normalised country codes, all 38 participating countries are typeset in bold.
[b] Area of country calculated from GISCO version 9 NUTS vector file rasterised at 250m using ETRS-89 LAEA projection and the European grid.

2.4 Merged coverages

Figures 2.16–2.18 illustrate the regions covered by the SPOT4, SPOT5, and IRS satellites respectively. The colour are used for coding the number of images overlapping a given area. The increasing order of overlap is mapped into the fundamental colour order and their combinations (white being used for more than 6 overlapping image). From Fig. 2.18, it can be seen that most of target territory is covered by this sensor. Noticeable exceptions include Iceland and all Atlantic Islands (beyond a series of small gaps). Figure 2.19 shows the degree of overlap when taking into account all three sensors. The maximum number of times a pixel is covered is equal to 16. Figure 2.20 shows a colour composition of the coverage by all 3 satellites.

Statistics for showing the percentage of country covered by each sensor and all sensors together are displayed in tables 2.6–2.7, the first for full data ROIs (i.e., DROIs) and the second for their reduction to the territory of the country the input image originates from (i.e., CROIs). Although per sensor statistics vary when using CROIs rather than DROIs, the same gaps are observed in both cases. Only 4 countries are not fully covered: Iceland (a few coastal pixels), Portugal (Porto Santo Island, Madeira), Sweden (Islet in the North Baltic sea), and UK (Saint Kilad Island). Note that the analysis is based on NUTS level 0 codes. Some islands such as Guernsey (a Channel Island) have separate codes so that their coverage is not secured and not taken into account in the table (in fact it happens that Guernsey Island is not covered).

Table 2.5 shows the maximal number of times a pixel is covered for each sensor and coverage. For example, in the first coverage there is a least one pixel that is covered by 10 different images (taking into account data ROIs). When considering all sensors and the combined coverages, there is a least one pixel that is covered 14 different images! Figure 2.15 shows the complete distributions.

Table 2.5: Maximum number of times a pixel is covered for each sensor and coverage (based on data/country regions of interest, the second value being given only when it differs from the first).

	SPOT4	SPOT5	IRS-P6	ALL
1st coverage	6/5	5	7	10/9
2nd coverage	8	5	6	12/10
merged coverages	8	6	11	14/13

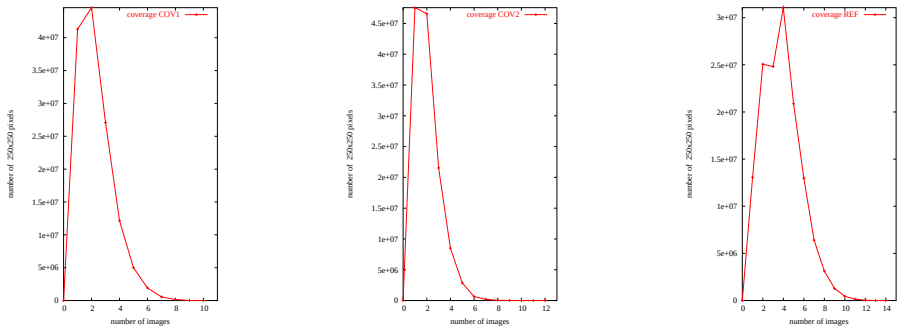


Figure 2.15: Number of pixels vs. number of images available for these pixels for each coverage (all sensors combined and using data ROIs). Coverage 12 means merged coverages.

Table 2.6: Combined coverage statistics based on data regions of interest. The code CS refers to the former union of Serbia (RS) and Montenegro (ME).

Code[a]	area[b] (km^2)	% with SP4	% with SP5	% with IL3	% with ALL	# missing pixels	Full coverage?
AL	28534	68.42	17.32	99.84	100.00	0	yes
AT	83930	64.68	28.37	100.00	100.00	0	yes
BA	51274	24.59	43.98	100.00	100.00	0	yes
BE	30664	39.19	59.02	100.00	100.00	0	yes
BG	110796	60.06	42.92	97.86	100.00	0	yes
CH	41288	60.32	36.86	100.00	100.00	0	yes
CS	102572	58.42	32.73	100.00	100.00	0	yes
CY	9247	0.00	8.39	100.00	100.00	0	yes
CZ	78866	44.83	36.70	99.73	100.00	0	yes
DE	357672	40.59	40.83	100.00	100.00	0	yes
DK	43360	80.47	60.67	97.53	100.00	0	yes
EE	45327	44.67	22.81	100.00	100.00	0	yes
ES	505984	12.37	20.48	98.07	100.00	0	yes
FI	337801	37.61	8.06	100.00	100.00	0	yes
FR	549167	43.85	26.23	98.82	100.00	0	yes
GR	132025	61.13	41.59	98.04	100.00	0	yes
HR	56631	16.35	38.86	100.00	100.00	0	yes
HU	93016	54.28	49.91	97.98	100.00	0	yes
IE	70175	74.12	47.94	100.00	100.00	0	yes
IS	102906	94.31	45.07	0.00	≈100.00	200	no
IT	301429	39.46	45.79	99.89	100.00	0	yes
LI	160	100.00	0.00	100.00	100.00	0	yes
LT	64892	61.17	55.37	98.01	100.00	0	yes
LU	2595	100.00	77.74	44.40	100.00	0	yes
LV	64603	65.66	13.14	100.00	100.00	0	yes
MK	25153	32.99	8.34	100.00	100.00	0	yes
MT	314	0.00	100.00	100.00	100.00	0	yes

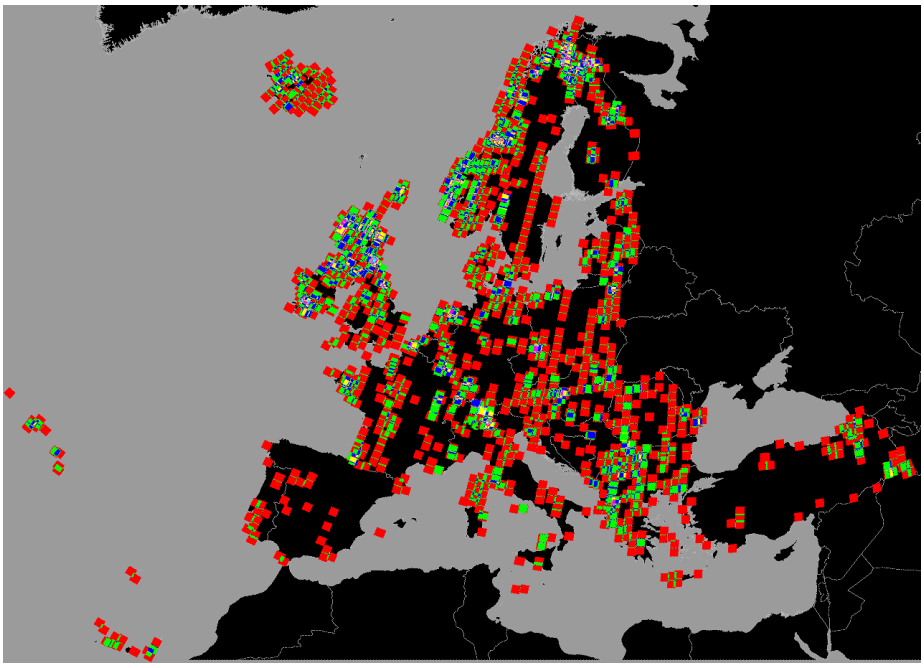


Figure 2.16: SPOT 4 footprints (combined coverages) based on data ROIs. The colours are used for coding the number n of images covering a given pixel: red ($n = 1$), green ($n = 2$), blue ($n = 3$), yellow ($n = 4$), magenta ($n = 5$), cyan ($n = 6$), and white $n \geq 7$. $n_{\max} = 9$.

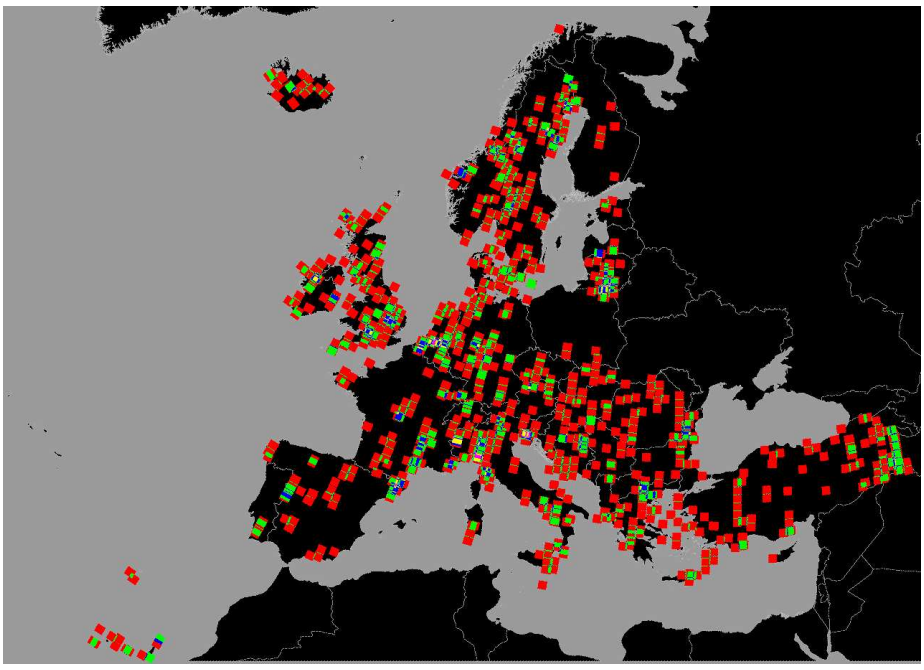


Figure 2.17: SPOT 5 footprints (combined coverages) based on data ROIs. The colours are used for coding the number n of images covering a given pixel: red ($n = 1$), green ($n = 2$), blue ($n = 3$), yellow ($n = 4$), magenta ($n = 5$), cyan ($n = 6$). $n_{\max} = 6$. Land (resp. sea) regions not covered are painted in black (resp. light grey)

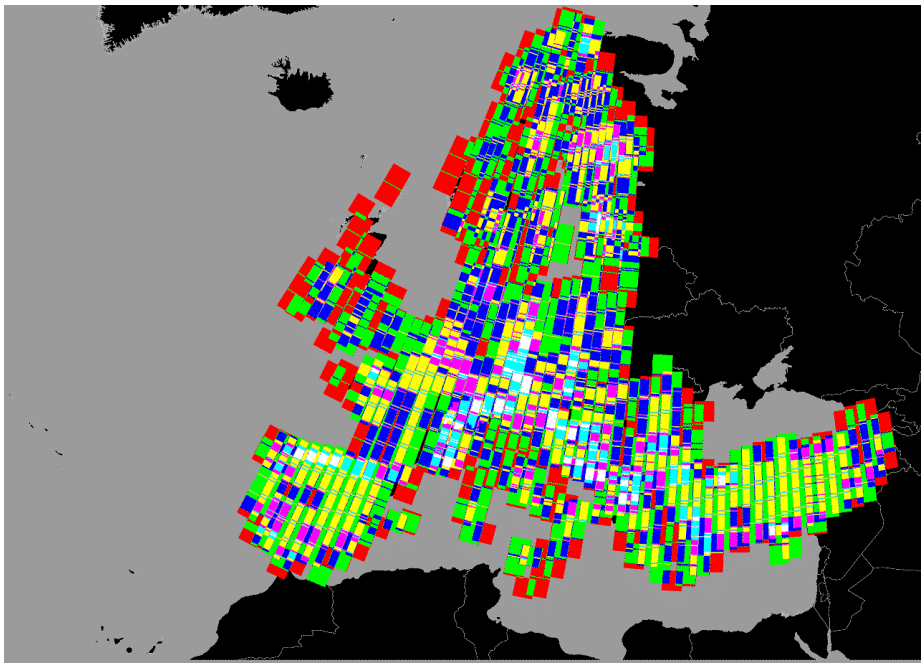


Figure 2.18: IRS-P6 LISS III-3 footprints (combined coverages) based on data ROIs. The colours are used for coding the number n of images covering a given pixel: red ($n = 1$), green ($n = 2$), blue ($n = 3$), yellow ($n = 4$), magenta ($n = 5$), cyan ($n = 6$), and white ($n \geq 7$). $n_{\max} = 11$. Land (resp. sea) regions not covered are painted in black (resp. light grey).

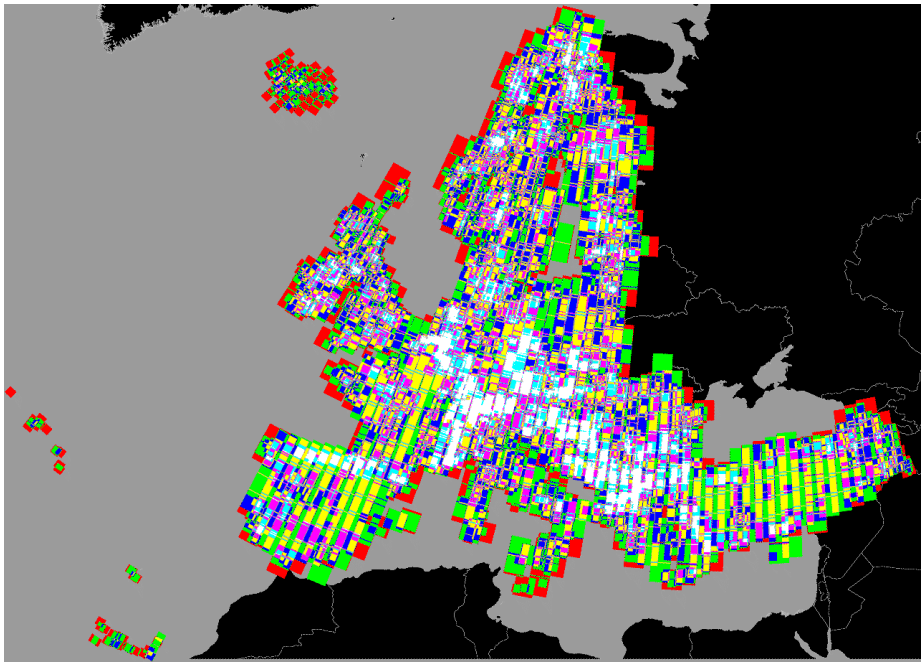


Figure 2.19: IMAGE 2006 combined coverages (all sensors together). The colours are used for coding the number n of images covering a given pixel: red ($n = 1$), green ($n = 2$), blue ($n = 3$), yellow ($n = 4$), magenta ($n = 5$), cyan ($n = 6$), and white ($n \geq 7$). $n_{\max} = 16$.

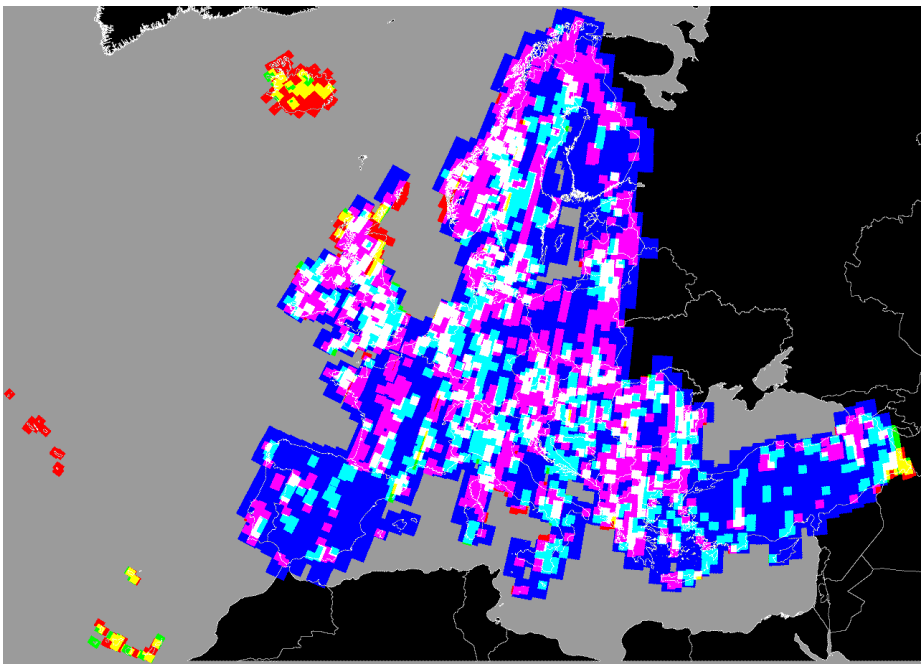


Figure 2.20: IMAGE 2006 combined coverages. The colours are used for coding the type of satellite covering a given area: red for SPOT 4, green for SPOT 5, blue for IRS. The combinations of these 3 fundamental colours indicate the type of combination, e.g., yellow for areas covered both by SPOT 4 and SPOT 5 and white for regions covered by all 3 satellites. For example, regions not covered by IRS appear in either red, green or yellow. Land (resp. sea) regions not covered are painted in black (resp. light grey).

Code[a]	area[b] (km^2)	% with SP4	% with SP5	% with IL3	% with ALL	# missing pixels	Full coverage?
NL	37357	82.22	76.52	98.81	100.00	0	yes
NO	323457	77.84	24.41	98.58	100.00	0	yes
PL	311894	46.56	9.80	99.21	100.00	0	yes
PT	92140	38.55	14.99	96.14	99.95	68500	no
RO	237938	56.74	40.32	99.03	100.00	0	yes
SE	449446	39.66	43.79	96.71	≈100.00	2200	no
SI	20272	82.40	41.18	97.62	100.00	0	yes
SK	49024	52.50	47.13	100.00	100.00	0	yes
TR	780122	17.40	28.25	98.68	100.00	0	yes
UK	244714	80.26	58.21	92.15	≈100.00	13200	no
ALL	5836765	44.01	32.25	96.77	≈100.00	84300.0	no

[a] Normalised country codes, all 38 participating countries (CS for RS ∪ ME).
[b] Area of country calculated from GISCO version 9 NUTS vector file rasterised at 250m using ETRS-89 LAEA projection and the European grid and within window displayed in Fig. 2.1.

Table 2.7: Combined coverage statistics based on country regions of interest. The code CS refers to the former union of Serbia (RS) and Montenegro (ME).

Code[a]	area[b] (km^2)	% with SP4	% with SP5	% with IL3	% with ALL	# missing pixels	Full coverage?
AL	28534	84.82	29.71	99.84	100.00	0	yes
AT	83930	80.91	39.02	100.00	100.00	0	yes
BA	51274	31.43	57.89	100.00	100.00	0	yes
BE	30664	49.77	69.08	100.00	100.00	0	yes
BG	110796	64.94	43.06	100.00	100.00	0	yes
CH	41288	69.58	40.76	100.00	100.00	0	yes
CS	102572	63.57	39.42	100.00	100.00	0	yes
CY	9247	0.00	8.39	100.00	100.00	0	yes
CZ	78866	54.67	37.51	99.88	100.00	0	yes
DE	357672	45.34	44.01	100.00	100.00	0	yes
DK	43360	80.59	60.67	99.43	100.00	0	yes
EE	45327	44.70	22.81	100.00	100.00	0	yes
ES	505984	13.75	21.59	98.07	100.00	0	yes
FI	337801	39.15	9.11	100.00	100.00	0	yes
FR	549167	44.47	26.59	98.84	100.00	0	yes
GR	132025	62.45	44.91	98.69	100.00	0	yes
HR	56631	26.93	58.39	100.00	100.00	0	yes
HU	93016	60.19	53.82	98.71	100.00	0	yes
IE	70175	80.47	49.00	100.00	100.00	0	yes
IS	102906	94.31	45.07	0.00	≈100.00	200	no
IT	301429	40.60	46.86	99.91	100.00	0	yes
LI	160	100.00	0.00	100.00	100.00	0	yes
LT	64892	70.15	55.67	100.00	100.00	0	yes

Code[a]	area[b] (km^2)	% with SP4	% with SP5	% with IL3	% with ALL	# missing pixels	Full coverage?
LU	2595	100.00	77.78	100.00	100.00	0	yes
LV	64603	69.14	14.83	100.00	100.00	0	yes
MK	25153	71.89	17.76	100.00	100.00	0	yes
MT	314	0.00	100.00	100.00	100.00	0	yes
NL	37357	83.64	76.53	100.00	100.00	0	yes
NO	323457	80.50	25.92	98.58	100.00	0	yes
PL	311894	49.42	10.56	100.00	100.00	0	yes
PT	92140	39.91	14.99	96.62	99.95	68500	no
RO	237938	59.52	42.84	99.61	100.00	0	yes
SE	449446	40.58	45.35	99.18	≈100.00	2200	no
SI	20272	88.75	41.65	100.00	100.00	0	yes
SK	49024	66.69	56.05	100.00	100.00	0	yes
TR	780122	18.06	28.53	98.68	100.00	0	yes
UK	244714	80.26	58.21	92.16	≈100.00	13200	no
ALL	5836765	46.79	34.10	97.18	≈100.00	84300.0	no

[a] Normalised country codes, all 38 participating countries (CS for RS ∪ ME).
[b] Area of country calculated from GISCO version 9 NUTS vector file rasterised at 250m using ETRS-89 LAEA projection and the European grid and within window displayed in Fig. 2.1.

Chapter 3

Cloud detection

Pierre Soille

Spatial Data Infrastructure Unit
Institute for Environment and Sustainability
DG Joint Research Centre, European Commission

3.1 Introduction

Because clouds mask the ground in the optical domain, cloud free images should be selected whenever possible. However, this is difficult if not impossible to achieve in some regions. For these regions, cloud free coverage is only possible thanks to multiple acquisitions. This explains why some areas are covered several times in the IMAGE-2006 coverages. The amount of cloud cover within the mosaic will be minimised because the developed morphological image compositing method takes cloud masks into account. Indeed, during mosaicing, cloud free image pieces are automatically selected. Hence, the automatic mosaicing procedure requires a cloud mask to be available for each input image.

This chapter details the methodology developed for automatically generating cloud masks. This methodology is based on the analysis of Top of Atmosphere (TOA) reflectance values rather than the input digital numbers. The conversion from input digital numbers to TOA reflectance values is based on Look-Up-Tables (LUTs) described in Sec. 3.2. The cloud detection method itself is detailed in Sec. 3.3 with its two main steps: unsupervised classification exploiting spectral attributes followed by a series of morphological transformations exploiting spatial attributes. Before concluding, results are presented and discussed in Sec. 3.4.

3.2 Digital numbers to TOA reflectance LUTs

The values of the pixels of any given input image are usually referred to as digital numbers (DNs). They can be converted to radiance values by applying

Table 3.1: Julian Day (J.D.) to Earth-Sun distance d in astronomical units. [Source: USGS, see http://eo1.usgs.gov/faq.php?id=25].

J.D.	d	J.D.	d	J.D.	d	J.D.	d	J.D.	d
1	.9832	74	0.9945	152	1.0140	227	1.0128	305	.9925
15	.9836	91	0.9993	166	1.0158	242	1.0092	319	.9892
32	.9853	106	1.0033	182	1.0167	258	1.0057	335	.9860
46	.9878	121	1.0076	196	1.0165	274	1.0011	349	.9843
60	.9909	135	1.0109	213	1.0149	288	0.9972	365	.9833

a linear transformation. However, because the parameters (i.e., gain and offset) of this linear transformation vary from one image to another, the input DNs originating from two different images are not necessarily comparable. Therefore, the application of fixed threshold levels to DNs is not robust. Even after conversion to radiance values, fixed threshold levels should not be applied because the radiance values depend themselves on the date and time of acquisition (varying Earth-Sun distance and sun elevation angle). This dependence can be suppressed (or at least mitigated) by converting the radiance values into reflectance values (i.e., proportion of the incident energy that is reflected). Fixed threshold levels can be applied to reflectance values.

Ideally, reflectance should be measured at the ground level. However, measurements are performed by the satellite at the TOA level. The conversion from TOA to ground reflectance values requires ancillary data about the atmospheric condition at the time of acquisition and/or user-interaction to select the pixels that are deemed to be the clearest. This is for instance the case of the haze optimised transform described in [47]. Given the available data and resources for the IMAGE-2006 Mosaic project, it was decided to use TOA reflectance values as a proxy for the desired ground reflectance values.

The TOA reflectance values are calculated using the gain, offset, acquisition date, sun elevation angle ϕ, and solar irradiance values retrieved from the header file of the corresponding band (see details in Sec. 1.4, page 13). The required formulæ can be found for example in [30] and are summarised hereafter. Denoting by x_i a digital number in the ith band, the corresponding TOA radiance L_i is calculated using the gain g_i and offset o_i values retrieved from the header file of the ith band:

$$L_i = (x_i - o_i)/g_i,$$

in $\mathrm{W}\times\mathrm{m}^{-2} \times \mathrm{sr}^{-1}\times\mu\mathrm{m}^{-1}$. The TOA reflectance ρ is then calculated by taking into account the sun zenith angle θ (complement of the sun elevation angle ϕ: $\theta = \pi/2 - \phi$) and the Earth-Sun distance d (in astronomical units) at the date of acquisition as well as the mean exoatmospheric solar irradiance E (in $\mathrm{W}\times\mathrm{m}^{-2}\times\mu\mathrm{m}^{-1}$) of band i:

$$\rho_i = \pi L_i d^2/(E_i \cos\theta).$$

The Earth-Sun distance can be interpolated from the values given in table 3.1. Alternatively, it can be calculated for any single Julian day[1] using astronomical

[1] Julian dates were calculated using the algorithm described in [16].

algorithms described in [25]. The reflectance is a unitless quantity whose values range from 0 (no incident energy reflected) to 1 (all incident energy is reflected).

In practice, the conversion from DNs to TOA reflectance values is achieved thanks to a TIFF image containing a unique line of 256 pixels with float values between 0.0 and 1.0. It corresponds to a LUT mapping the DNs to the TOA reflectance values. For example, the TOA value of the digital number with value 0 is given by the value of the first pixel of this LUT. By doing so, the above formulæ need to be computed only once for each 256 possible DN value rather than for each pixel.

3.3 Methodology

The Automatic Cloud Cover Assesment (ACCA) put forward in [18] is a proven method for detecting clouds with Landsat imagery. ACCA is an unsupervised classifier for clouds that takes advantage of known properties of clouds, snow, bright soil, vegetation, and water. A detailed analysis of the reliability of this algorithm as well as its failure modes (mainly semi-transparent clouds such as of cirrus and cloud edges) is reported in [19]. Beyond the spectral bands available for SPOT-4 HRVIR, SPOT-5 HRG, and IRS-LISS III imagery, ACCA exploits the measured reflectance values in the blue, mid infrared in the range 2.08–2.35 µm, and thermal bands. Unfortunately, these latter 3 bands are not available for these sensors. To solve this problem, the proposed methodology relies on a modified ACCA algorithm (Sec. 3.3.1) followed by a series of morphological transformations exploiting spatial attributes (Sec. 3.3.2). All stored data is summarised in Sec. 3.3.3.

3.3.1 First step: modified ACCA algorithm

For the purpose of IMAGE-2006 mosaicing, the automatic cloud cover assessment (ACCA) algorithm developed in [18, 19] for Landsat images was adapted to SPOT and IRS images. Given that only 4 of the 7 bands of Landsat TM imagery are available for SPOT0-4 HRVIR, SPOT-5 HRG, and IRS-LISS III imagery (see above), only ACCA filters 1–2 and 5–7 of the so-called 'pass one' [18] can be retained without any modification. The other filters have been replaced by ad hoc procedures. ACCA filter 3 (originally temperature threshold based on thermal band) was replaced by a threshold on the arithmetic difference between band 3 and band 2. ACCA Filter 4 (originally Landsat band 5/band 6 composite) was replaced by an intensity threshold on band 1. ACCA filter 8 has been discarded as well as ACCA 'pass two' processing because they require the availability of the thermal channel [18]. A brief description of each filter is given hereafter[2]:

- Filter 1: brightness threshold. All pixels in band 2 whose reflectance is below 0.08 are classified as non-cloud [18].

- Filter 2: normalised difference snow index (NDSI). This index is defined

[2]Due to the absence of the blue channel for SPOT-4 HRVIR, SPOT-5 HRG, and IRS-LISS III images, the band number indices given in this report need to be incremented by one when considering Landsat data.

as follows [17]:

$$\text{NDSI} = (\text{band } 1 - \text{band } 4)/(\text{band } 1 + \text{band } 4).$$

All pixels whose NDSI exceeds 0.7 are classified as non-cloud [18].

- Filter 3: difference between band 3 and band 2. All pixels whose reflectance in band 3 is not at least 0.05 greater than in band 2 are classified as non-cloud.
- Filter 4: intensity threshold on band 1 only. All pixels whose band 1 reflectance is less than 0.1 are classified as non-cloud.
- Filter 5: band 3/band 2 ratio. All pixels whose ratio exceeds 2.0 are classified as non-cloud [18].
- Filter 6: band 3/band 1 ratio. All pixels whose ratio exceeds 2.0 are classified as non-cloud [18].
- Filter 7: band 3/band 4 ratio. All pixels whose ratio exceeds 1.0 are classified as non-cloud [18].

All potential cloud pixels in the image mask corresponding to the ith filter are set to 2^{i-1}, the non-cloud pixels being set to 0. This enables the combination of all 7 individual masks into a unique image by simply computing the bitwise union operation between all 7 individual masks. Therefore, in this resulting combined image, the maximum value of a pixel equals 127 (all 7 filters returning potential cloud). Examples are displayed in the middle image column of Fig. 3.1 (the generation of the binary cloud masks displayed in the right column is detailed in the next section). The input images have been rendered by mapping the bands 4, 3, and 1 to the red, green, and blue channels (false colour composition).

3.3.2 Second step: spatial context

Unfortunately, due to the lack of a thermal band, those pixels of the `IM-ACCA` images that return potential clouds for all 7 filters (i.e., pixels with value 127 in `IM-ACCA`) lead to mainly true positive detection but contain many false negative errors. Therefore, these pixels are merely considered as markers indicating the presence of clouds with a high degree of confidence. These markers are then used as seeds for initiating a region growing process expanding the marker set until the full cloudy regions are reached. More precisely, the binary mask of clouds is generated by considering the morphological reconstruction by dilation [33] of the pixels with value 127 (i.e., potential clouds for all 7 filters) using as geodesic mask these pixels plus those with value 79 (all but filters 5 and 6 returning cloud), 95 (all but filter 6 returning cloud), and 111 (all but filter 5 returning cloud). The resulting mask is then cleaned by filling all its holes and then suppressing all connected components not containing at least one cloud block of 4×4 pixels. Note that, with the proposed methodology, clouds need to contain at least one pixel with value 127 in the ACCA image because otherwise the region growing process cannot be initiated.

Examples showing the input images, the images combining the output of all 7 individual filters, and the resulting binary cloud masks are given in Fig. 3.1.

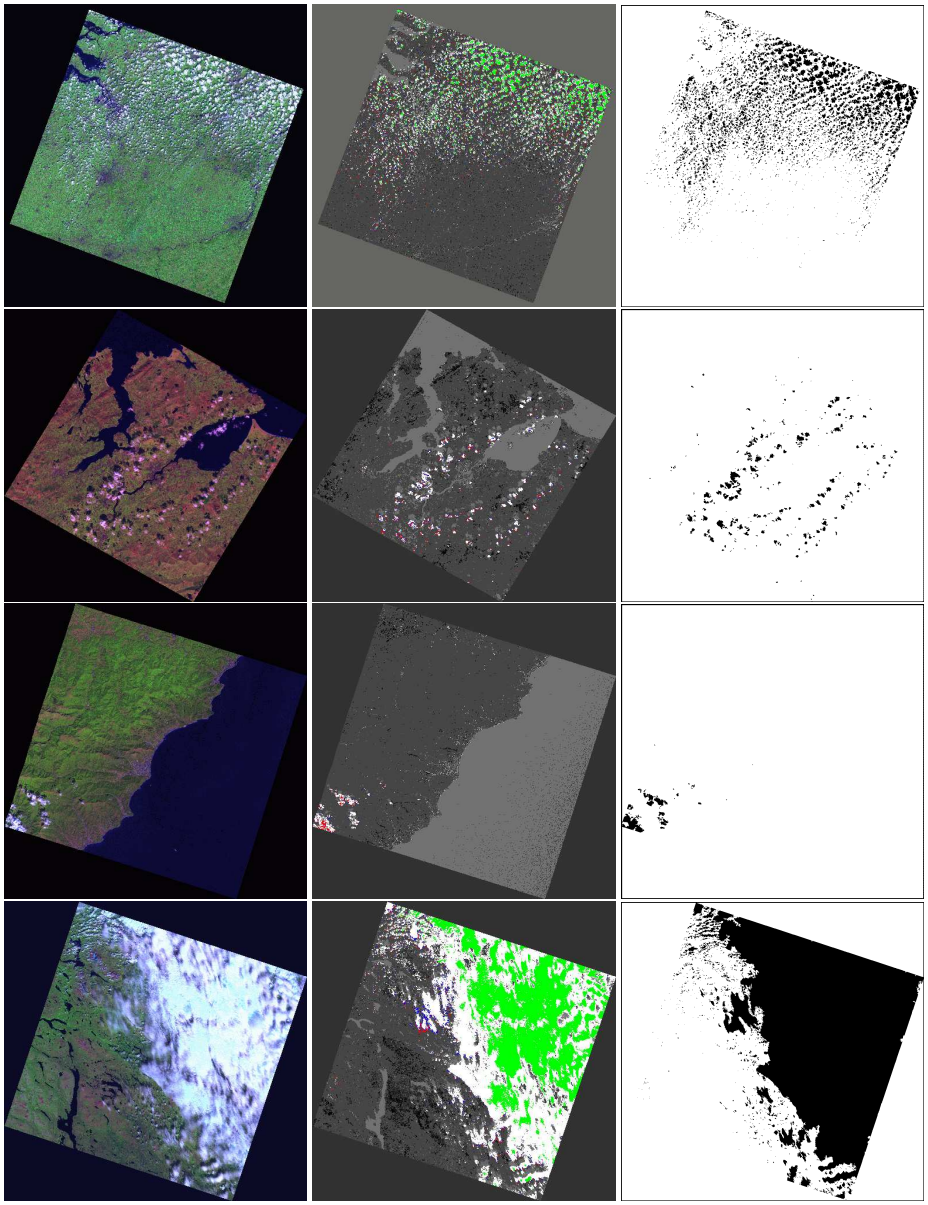


Figure 3.1: IMAGE-2006 cloud detection. Left column: input images. Middle column: output of modified ACCA filters coded on 7 bit planes with white for code 127 (all filters returning potential clouds), red for code 79 (all but filters 5 and 6 returning cloud), green for code 95 (all but filter 6 returning cloud), and blue for code 111 (all but filter 5 returning cloud). Right column: Resulting binary cloud masks. Input images (from top to bottom, all from coverage 1): `20060710-1051 IL3 BE 3955931097-CB`, `20060404-1149 SP4 IE 3225036819-CB`, `20060830-1035 SP5 IT 4179723343-DA`, and `20060715 IL3 NO 4469843797-AB`.

3.3.3 Stored data

All relevant data related to cloud detection is stored in the ancillary directories `2006_COV1_ANC`, `2006_COV2_ANC`, and `2006_REF_ANC`:

- DN to TOA LUTs. For the reference coverage, their type and identifier fields are set to the following values: `IP-L10C` for band 1, `IP-L20C` for band 2, etc. For the individual coverages, the `0C` characters are replaced by the appropriate country code `XX`: `IP-L1XX`, etc.;
- ACCA outputs with bitwise coding of each filter: For the reference coverage, these results stored in the image file using `IM-ACCA` as type/identifier field (plus the country code for images originating from the individual coverages: `IM-ACCAXX`);
- binary cloud masks: The type and identifier field of the binary cloud masks is set to `IM-CLDS` in the case of the reference coverage. For the individual coverages, the `DS` letters are replaced by the two letter country code: `IM-CLXX`.

All this data occupies approximately 16 Gbytes on disk.

3.4 Results

The produced cloud masks are evaluated in Sec. 3.4.1. Success and failure modes of the proposed methodology are also reported and illustrated. Summary statistics are presented in Sec. 3.4.2.

3.4.1 Cloud masks

Overall, for the purpose of minimising cloud cover when mosaicing, the quality of the produced cloud masks is sufficient. A gallery of 27 examples of satisfactory cloud masks in a range of landscapes and latitudes is displayed in Fig. 3.2. Nevertheless, given the absence of a thermal band, it is not possible to avoid false positive and false negative errors when detecting clouds from SPOT-4 HRVIR/SPOT-5 HRG and IRS-LISS III images with the procedure described in Sec. 3.3. A gallery illustrating cases where the algorithm is failing to detect clouds or wrongly detecting non-cloud objects as clouds is displayed in Fig. 3.3. Most false positive errors are due to ice and snow fields detected as clouds. Fortunately, this type of error is not an issue in view of creating a mosaic because it is usually desirable to minimise snow/ice covered areas. For instance, a 30th of December scene with snow will be automatically discarded (see example in Fig. 3.6). A series of false negative errors may also occur. This is the case for some bare soil fields, roofs of large buildings, and a few beaches detected as clouds (see gallery). Furthermore, semi-transparent clouds and thin clouds over the sea are also missed. This latter limitation is also reported for the original ACCA procedure on Landsat imagery [19]. Note that the detection of clouds over the sea is not critical for mosaicing purposes since, although undesirable, they are not masking out the terrain.

All images of Figs. 3.2–3.3 were extracted from a mosaic image combining all available imagery of the first coverage by performing the point-wise maximum

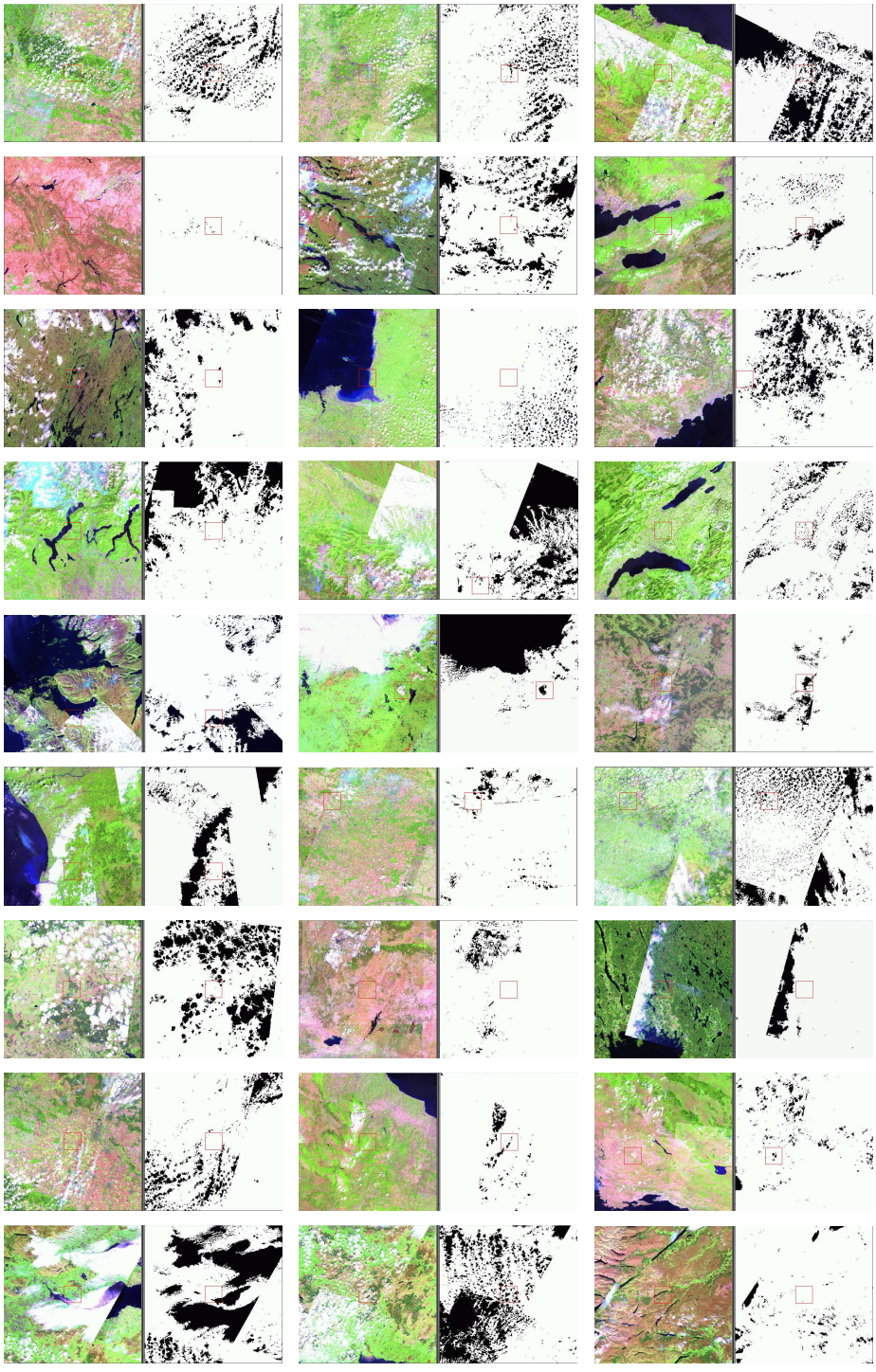


Figure 3.2: Gallery of satisfactory cloud masks.

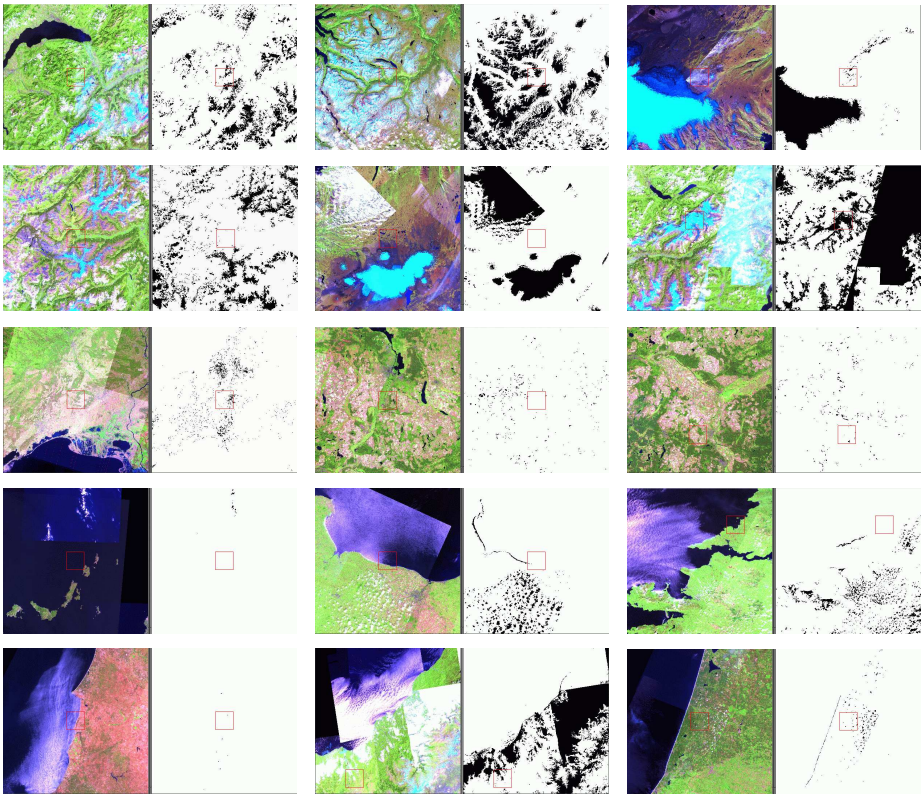


Figure 3.3: Gallery of unsatisfactory cloud masks (false positive errors and, to a lesser extent, false negative errors). Top two rows: ice and snow detected together with clouds. Third row: bright bare agricultural fields detected as clouds. Two bottom rows: undetected clouds over the sea (and beaches/breaking waves detected as clouds, last image).

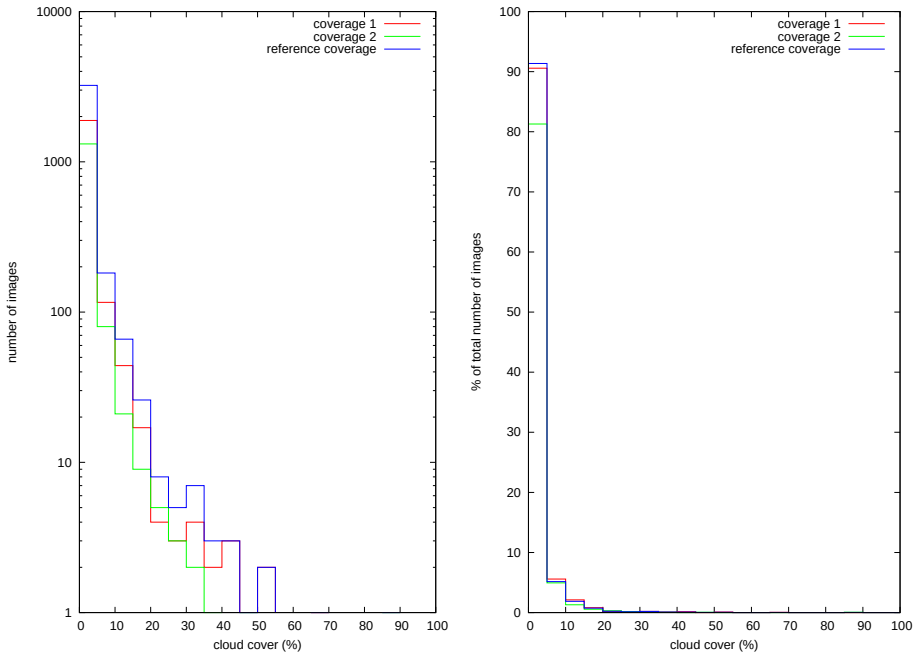


Figure 3.4: Frequency distribution of the cloud cover (in percentage of full data ROI) for all scenes of coverages 1 and 2 as well as the reference coverage. Left: in terms of number of images (logarithmic scale for y-axis). Right: normalised against the total number of images of each coverage. The histograms were produced using a bin width of 5% cloud cover.

operation between all images. It follow that all clouds are revealed in this mosaic because clouds are reflecting more than land cover classes in the considered bands. Similarly, a unique cloud mask was obtained by considering the union of all generated cloud masks.

3.4.2 Summary measurements

Once the cloud masks are available, summary statistics can be produced for each coverage. For example, Fig. 3.4 shows the histograms of the cloud percentage for all scenes of each coverage using a bin width of 5% cloud cover, the percentage being calculated on the basis of DROIs. The first histogram indicates the number of images and the second the percentage of total number of images for each bin. Figure 3.5 shows the union of all cloud masks obtained for coverages 1 and 2 respectively. For instance, in the latter case, it can be seen that more than 90% of the images of the reference coverage contain less than 5% of clouds. There is a total of 123 images whose calculated cloud coverage percentage exceeds 10% of their DROIs (not taking into account that clouds may fall in the sea only so that percentages may improve for some coastal scenes if only land were taken into account). Table 3.2 lists the identifiers of the 31 images of the reference coverage whose cloud cover is greater than 20% of their data ROI. The 4 scenes with cloud masks covering more than 50% of their full DROI are

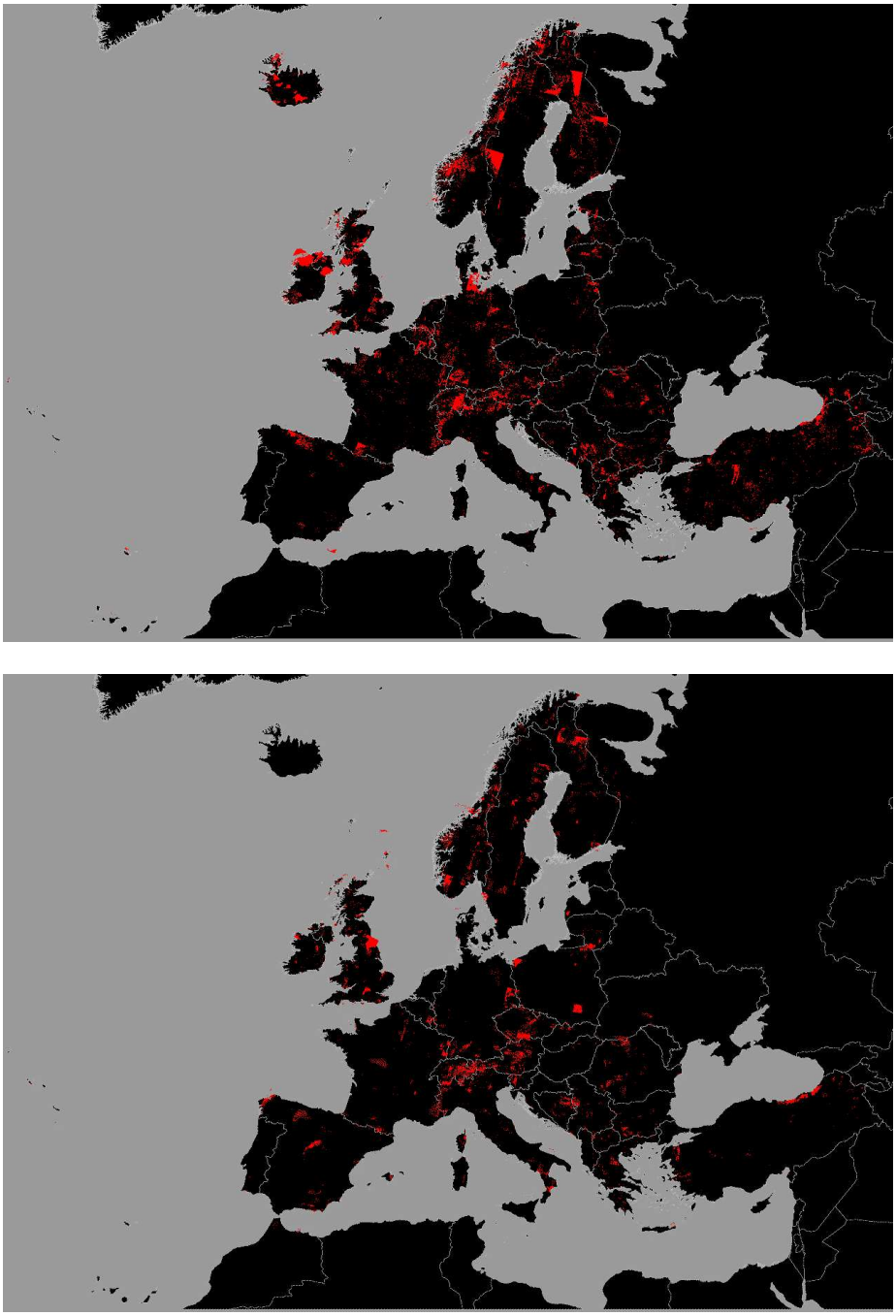


Figure 3.5: Cloud masks. Top: composition of the cloud masks of all 2,080 scenes of the first coverage. Bottom: composition of the cloud masks of all 1,619 scenes of the second coverage.

Table 3.2: Unique identifiers of the images of the reference coverage leading to a cloud cover greater than 20% of their data ROI (31 files out of a total of 3,533 files).

tile	date-time	sensor	centre	cloud (%)
43-B	20060426-1012	IL3	4635831544-CB	20
42-C	20060504-1033	SP4	4167128198-BC	20
43-A	20060916-1034	IL3	4370734847-DB	21
54-C	20060806-0944	IL3	5236446246-DA	22
33-C	20070529-1213	SP4	3060436246-DB	22
44-A	20061101-1050	SP4	4276940069-AA	22
44-A	20050630-1103	IL3	4247043769-BC	23
44-A	20070701-1103	SP5	4150943404-CC	24
44-D	20050702-1019	IL3	4908149296-AB	26
42-C	20060515-1022	SP4	4145927675-AB	27
44-D	20060502-1037	SP5	4771146249-AB	27
44-A	20070822-1134	SP4	4137742927-BA	28
62-B	20050730-0801	SP4	6867024961-DB	28
54-A	20060415-0954	SP4	5219142617-DB	30
25-B	20060909-1249	SP4	2917351133-CC	32
33-D	20070502-1125	IL3	3550335486-BA	32
33-C	20060817-1154	SP4	3217937133-AB	32
44-D	20070919-1022	SP5	4501346948-CA	32
44-D	20060809-1104	SP4	4538747441-CB	33
42-C	20060902-1028	IL3	4246726010-DC	34
33-C	20060908-1131	SP4	3443436337-AC	37
24-D	20070802-1300	SP4	2883549742-CB	37
42-C	20060610-1022	SP4	4133925259-BB	38
54-C	20060805-1005	IL3	5048649609-BB	43
42-C	20060610-1022	SP4	4222826061-AC	44
42-C	20060610-1022	SP4	4236326591-DB	44
42-C	20051230-1037	SP4	4244025564-BB	46
33-C	20050712-1156	IL3	3083036512-CB	51
44-A	20060715-1044	IL3	4469843797-AB	53
34-C	20060819-1253	SP4	3070748369-BB	69
53-A	20061230-0946	SP5	5092530859-DB	85

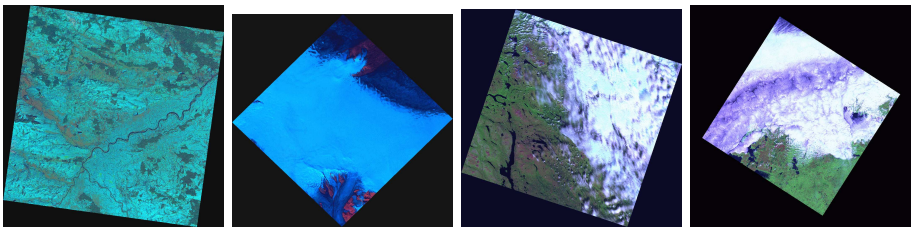


Figure 3.6: The 4 scenes with cloud masks deemed to be greater than 50% of their DROI. The names of these scenes are listed in the last four lines of table 3.2. They are displayed here in order of decreasing cloud cover. Note that the first scene is a 30th of December scene in the South-East of Poland with snow covered ground while the second scene shows part of the Vatnajökull glacier in Iceland.

Table 3.3: Clouds remaining within the territory of the participating countries when considering available imagery for each coverage in percentage of DROIs and CROIs.

	DROI			CROI		
	COV1	COV2	REF	COV1	COV2	REF
# Kpixels	25909	21660	10167	60322	44174	19836
Area (sqkm)	16193	13537	6354	37701	27608	12397
Area (%)	0.27	0.24	0.1	0.64	0.49	0.21

displayed in Fig. 3.6.

Fortunately, most cloudy areas are covered by more than one image so that most clouds will not appear in the produced mosaics. Indeed, when clouds are present in a region, the mosaicing algorithm automatically selects an image with no clouds if available. For instance, table 3.3 indicates the area of the regions that are always covered by clouds when considering all available imagery. For creating this table, measurements have been restricted to the territory of the participating countries. This table reveals that by substituting CROIs with DROIs, the permanent cloud cover is almost halved in all cases. For example, in the first coverage it would go from 0.64% to 0.27%. However, the quality of the orthorectification is only guaranteed within the country based ROIs. Permanent cloud area measurements performed at the country level are presented in table 3.4. Again, it can be observed that the use of DROIs rather than CROIs would often lead to a drastic decrease of the percentage of permanent cloud cover. Because snow and ice is usually detected as clouds (most common false positive error), all countries having large areas permanently covered with snow or ice show higher cloud percentage cover for all available imagery. This is the case Austria (AT), Switzerland (CH), Iceland (IS). Additional countries with remaining cloud cover larger than 1/1000 in the first coverage are Ireland (IE), Italy (IT), FYROM (MK), Malta (MT), Sweden (SE), Turkey (TR), and the United Kingdom (UK). The comprehensive table indicating permanent cloud cover statistics for all countries at least partly covered by the IMAGE-2006

Table 3.4: Clouds remaining within each participating countries when considering available imagery for each coverage in percentage of DROIs and CROIs. In this table, percentages are relative to the portion of the territory covered by available imagery rather than the whole territory).

	DROI			CROI		
	COV1	COV2	REF	COV1	COV2	REF
AL	0.03	0.13	0.00	0.21	1.02	0.07
AT	1.23	0.29	0.05	2.27	2.23	0.75
BA	0.00	0.08	0.00	0.43	1.06	0.01
BE	0.00	0.00	0.00	1.12	0.70	0.02
BG	0.03	0.48	0.00	0.28	0.75	0.00
CH	0.80	0.22	0.07	5.31	4.81	0.70
CS	0.04	0.03	0.00	0.70	0.17	0.00
CY	0.03	0.00	0.00	0.03	0.00	0.00
CZ	0.02	0.02	0.00	0.33	2.08	0.09
DE	0.06	0.02	0.00	0.29	0.28	0.00
DK	0.04	0.00	0.00	0.50	0.00	0.00
EE	0.04	0.02	0.00	0.04	0.02	0.00
ES	0.03	0.22	0.00	0.05	0.23	0.00
FI	0.03	0.30	0.00	0.20	0.31	0.02
FR	0.08	0.10	0.00	0.37	0.28	0.01
GR	0.03	0.30	0.00	0.36	0.44	0.00
HR	0.04	0.03	0.00	0.36	0.12	0.01
HU	0.01	0.00	0.00	0.76	0.02	0.00
IE	0.26	0.23	0.00	1.78	0.73	0.00
IS	4.92	-	4.92	4.92	-	4.92
IT	0.32	0.13	0.01	0.90	0.27	0.04
LI	0.00	0.00	0.00	0.46	0.15	0.15
LT	0.07	0.01	0.00	0.17	0.08	0.00
LU	0.00	0.00	0.00	1.20	2.38	0.73
LV	0.10	0.01	0.00	0.45	0.09	0.02
MK	0.11	0.00	0.00	1.24	0.66	0.14
MT	0.89	0.03	0.00	0.89	0.03	0.00
NL	0.00	0.03	0.00	0.46	0.07	0.00
NO	0.74	1.61	0.18	1.23	1.72	0.36
PL	0.04	0.01	0.00	0.17	0.47	0.01
PT	0.03	0.19	0.03	0.03	0.19	0.03
RO	0.06	0.21	0.02	0.08	0.40	0.02
SE	0.11	0.28	0.05	1.46	0.48	0.95
SI	0.01	0.11	0.00	2.30	3.43	0.32
SK	0.01	0.03	0.00	0.19	0.53	0.01
TR	0.32	0.18	0.02	0.33	0.25	0.02
UK	0.56	0.33	0.00	0.59	0.34	0.00

imagery is given in appendix D.

3.4.3 On cloud shadows

Beyond clouds, the occurrence of cloud shadows should be minimised in the produced mosaics. Rather than relying on complex spectral rules for detecting shadows, the generated masks are expanded on the fly during mosaicing so as to cover potential shadows. This expansion is achieved by performing a morphological dilation by a line segment [39, 37]. The length of the line segment is determined by the sun elevation angle and by assuming that clouds lie 2.5km above the ground. This assumption is valid for the most common type of clouds cluttering the IMAGE-2006 imagery, i.e., fair weather cumulus clouds (puffy cotton balls floating in the sky). Note that the directional dilation produces a continuous mask going from the detected cloud and reaching the estimated position of its shadow. This secures that the same image will be used to patch the cloud and its shadow. Consequently, the resulting mosaic has a higher spatial coherence than those authorising one to switch from one image to another when masking out a cloud and its shadow separately.

3.5 Conclusion

Although the algorithm described in this section was deemed sufficient for the purpose of creating cloud free mosaic, higher quality masks could probably be obtained using more elaborate algorithms in the future. Note that there exist very few publications detailing algorithms for cloud detection on SPOT and IRS imagery. Apart from basic procedures based on a unique intensity threshold level to identify the best cloud-free pixels among the pixels from the multiple images of the same region [24], most algorithms use very complex procedures based on neural networks [9, 20] so that they are not easily repeatable. Recently, a procedure based on Markov random fields was also put forward [22].

A simpler and potentially more effective method could exploit the availability of multiple imagery over the same areas to detect clouds thanks to change detection methods (for instance, clouds usually appear as well defined clusters in the bivariate histogram computed for the same bands within the overlapping region of a pair of images, see examples of scattergrams on page 72). Another promising approach could be based on the spectral rule based preliminary mapping procedures detailed in [11].

Chapter 4

Relative geometric and radiometric accuracy

Pierre Soille and Jacopo Grazzini

Spatial Data Infrastructure Unit
Institute for Environment and Sustainability
DG Joint Research Centre, European Commission

4.1 Introduction

Within their domain of overlap, two images may differ in both geometry and radiometry. Consequently, when they are mosaiced, these differences may reveal the position of the seam line even if they follow salient image structures such as roads, streams, or field boundaries. A pair of overlapping images is said to be *consistent* if they are in agreement to one another in both geometry and radiometry. Geometric consistency measurements are derived from normalised cross-correlation computations in the spatial domain. Radiometric consistency measurements are derived by estimating the correlation coefficients and parameters of the linear regressions mapping the reflectance values of an image to those of its overlapping image.

The input 3,699 IMAGE-2006 images lead to 29,447 pairs of overlapping images when considering data region of interest (DROIs). This number decreases to 26,119 when considering country based regions of interest (CROIs). The number of overlapping pairs between all possible sensor combinations is given in table 4.1.

Contrary to consistency measurements, *accuracy* measurements require comparisons with a reference data set that can be considered as the ground truth. Geometric accuracy with respect to the reference used for orthorectification (IMAGE-2000 [28] and USGS Land Cover reference [42] datasets, the latter for all countries that were not participating to IMAGE-2000) is detailed elsewhere [27].

Table 4.1: Number of between sensor overlapping pairs based on DROIs (left) and CROIs (right). Overlapping pairs where the search for control points lead to less than 7 points have been discarded.

	SP4	SP5	IL3
SP4	3608	2668	7888
SP5	2668	1296	4812
IL3	7888	4812	6996

	SP4	SP5	IL3
SP4	3232	2366	6460
SP5	2366	1176	3942
IL3	6460	3942	5923

This chapter is organised as follows. The methods used for assessing the geometric and radiometric consistency are described in Secs. 4.2 and 4.3 respectively. The type of stored data is detailed in Sec **??**. Summary statistics for overlapping image pairs originating from the same or different sensors as well as those from the same or different countries are presented in Sec. 4.5. Concluding remarks are the subject of Sec. 4.6.

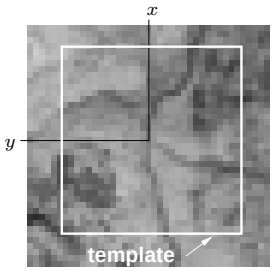
4.2 Geometric consistency

Geometric consistency is determined using normalised cross-correlation [12, 23]. In mathematical terms, the normalised cross-correlation γ function between an image f and a template t is defined as follows:

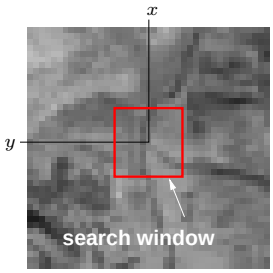
$$\gamma(u,v) = \frac{\sum_{x,y}\left[f(x,y) - \overline{f}_{u,v}\right]\left[t(x-u,y-v) - \overline{t}\right]}{\sqrt{\sum_{x,y}\left[f(x,y) - \overline{f}_{u,v}\right]^2}\sqrt{\sum_{x,y}\left[t(x-u,y-v) - \overline{t}\right]^2}}, \qquad (4.1)$$

where the sum is carried over the window containing the template t centred at (u,v), $\overline{f}_{u,v}$ denotes the mean value of $f(x,y)$ within the domain of the template t shifted to (u,v), and $\overline{t}$ denotes the mean value of the template. That is, $\gamma(u,v)$ gives a measure of the degree of similarity between an image and a template centred at coordinates (u,v). The measure is normalised in the sense that the observed intensity values are modified by subtracting their mean value and dividing them by their standard deviation. Consequently, the normalised cross-correlation is invariant to changes in image amplitude such as those caused by changing illumination conditions. The range of the correlation function is equal to the interval $[-1,1]$, a value of 0 being obtained for uncorrelated windows.

In practice (see Fig. 4.1), given a pair of overlapping images, we arbitrarily call the first image the *anchor* image and the second the *slave* image. A template of fixed size is created by cropping the anchor image at a given position corresponding to a node of a sampling grid. The normalised cross-correlation function is computed in the spatial domain within a search window of fixed size located at the same position in the slave image. The vector separating this position from the position at which the maximum value of the normalised cross-correlation function occurs is used as a local estimation of the translation between the anchor and slave images. The subpixel position and value of the maximum is calculated by interpolating the correlation function. This interpolation is achieved by determining the paraboloid passing through the 5 points corresponding to the point maximising the discrete correlation function and its

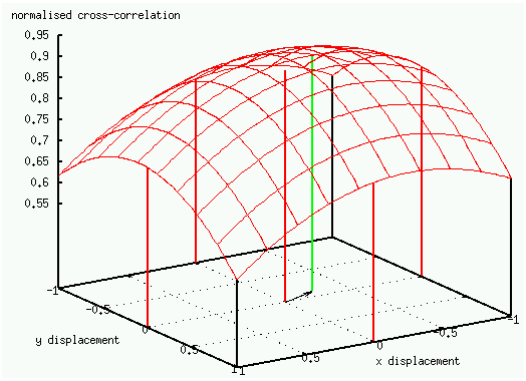


(a) Anchor image



(b) Slave image

(c) Cross correlation function γ and its 3-D representation



(d) Subpixel interpolation of γ around its peak

Figure 4.1: Geometric consistency assessment between an overlapping image pair: example of cross correlation calculations at a grid node of coordinates (x, y). See text for details.

four direct neighbours, see points materialised by the five vertical red line segments in Fig. 4.1d. The position at which the maximum of this paraboloid occurs (see green line segment in Fig. 4.1d) is used as subpixel estimation of the actual maximum of the correlation function. The maximum interpolated value corresponds to the output of the paraboloid function at this position (the upper bound being set to 1.0). This subpixel registration method is known as correlation interpolation by a quadratic estimator [41]. With this method, the accuracy of the lateral translation measurements reduces from ± 0.5 to ± 0.1 pixel (worst case). The actual accuracy depends on how well the correlation function around its peak approximates a quadratic function [15].

Only those points whose cross-correlation function lead to a clear maximum with isotropic behaviour are retained. More precisely, the following steps are considered for each pair of overlapping images:

- determine the domain of overlap between anchor and slave;
- for each node of a pan-European grid of width of 40 pixels (i.e., the coordinates of all nodes are equal to a multiple of 1000m) falling within the domain of overlap and without clouds[1], determine the position and magnitude of the maximum of the normalised cross-correlation using a template of width equal to 31 pixels and a search width in the slave image equal to 15 pixels. That is, it is expected that the maximum horizontal and vertical displacements does not exceed 7 pixels and that a template pf 31 pixels is large enough to ensure that high correlation values will only

[1] Clouds were detected using the method described in [35].

be obtained for actually matching points.

- select all points whose maximum normalised cross-correlation is greater than or equal to 0.75 and whose aspect ratio (ratio of major to minor axis) is less than or equal to 1.1;
- calculate the x and y displacements for each selected point;
- output in an ASCII file the mean, the root mean squared error (RMSE), and the standard deviation of all calculated displacements.

Given the sheer amount of image pairs to test, the computations were performed on a distributed system so that up to 24 pairs of overlapping images were processed in parallel.

4.3 Radiometric consistency

Radiometric consistency between all pairs of overlapping images is evaluated by comparing either the digital numbers or the top of atmosphere values of each pair of intersecting images within their domain of overlap not contaminated by clouds. This is achieved by determining the coefficients of the linear transformation mapping the values of the slave image to that of the anchor image (method sometimes called 'simple regression' [21, 46]). These coefficients are calculated for each band by performing a linear regression (one per band). In mathematical terms, denoting by x (resp. y) the top of atmosphere reflectance values (or the input digital numbers) in the anchor (resp. slave) images, the following linear transformation is estimated:

$$y^\star = ax + b.$$

The coefficients of this linear transformation are determined by minimising the sum of the square of the differences between the observed and estimated values: $\sum_i (y_i^\star - y_i)$. This leads to coefficients expressed in terms of the mean μ and standard deviation σ values:

$$\begin{aligned} a &= \sigma_{XY}/\sigma_X, \\ b &= \mu_Y - a\mu_X. \end{aligned}$$

The correlation coefficient ρ_{XY} indicates the strength and direction of the calculated linear relationship between the slave and anchor intensity values:

$$\rho_{XY} = E\big[(X - \mu_X)(Y - \mu_Y)\big]/(\sigma_X \sigma_Y). \tag{4.2}$$

Note the similarity of this equation with the normalised cross-correlation function at the basis of the geometric consistency measurements (Eq. 4.1). Finally, the residual variance of the errors (referred to as err in the generated file) is also computed:

$$\text{err} = \sum_i (x_i - y_i^\star)^2/n = (1 - \rho_{XY}^2)\sigma_Y^2,$$

with n equal to the number of pixels in the ROI without clouds. Perfectly matching overlapping image pairs have unitary slopes (a), zero offset (b), unitary correlation coefficients ρ_{XY}, and zero residual variance err.

```
ANCHOR 2006_COV1/42-A/20060615-1027_SP4_IP-B4FR_4130223241-DB.tif
SLAVE  2006_COV1/42-A/20070726-1018_SP4_IP-B4FR_4083323430-AB.tif

NPIX_IN_ROI 947443
NPIX_IN_ROI_WITHOUT_CLOUDS 884756
PERCENTAGE_WITHOUT_CLOUDS 0.933836

RELATIVE GEOMETRIC CONSISTENCY:
TEMPLATE_WIDTH 31
SEARCH_WIDTH 15
GRID_WIDTH 40
PIXELS_WITH_NCC_CALCULATED 975.0
PIXELS_WITH_ABS(NCC)_GEQ_0.750000 305.0
PIXELS_WITH_ABS(NCC)_GEQ_0.750000_AND_ELLIPSIS_ASPECT_RATIO_LEQ_1.100000 293.0
X_MEAN_(m) 17.568
Y_MEAN_(m) -52.159
X_RMSE_(m) 18.834
Y_RMSE_(m) 52.204
X_STD_(m) 6.788
Y_STD_(m) 2.179

RELATIVE RADIOMETRIC CONSISTENCY:
REG_PARAM(dn) [a b corr err]
  1:  0.856  16.774  0.688  696.138
  2:  0.568  36.711  0.530  383.561
  3:  0.976  34.118  0.749  1027.421
  4:  0.675  31.181  0.729  517.678
REG_PARAM(toa) [a b corr err]
  1:  0.897  8.199  0.687  164.691
  2:  0.592  24.698  0.529  168.559
  3:  0.686  7.535  0.749  50.235
  4:  0.713  8.030  0.729  34.046
```

Figure 4.2: Output measurements of geometric and radiometric consistency checks. The rather high geometric and radiometric deviations between this pair of images is highlighted on Fig. 4.3.

4.4 Stored data

Output measurements for both geometric and radiometric analysis are stored in an ASCII file whose name contains the names of the anchor and slave images. An example of output measurements produced for an overlapping image pair is displayed in Fig. 4.2. This file highlights that the relative displacement between the two input scenes exceeds 2 pixels in the y direction and is close to 1 pixel in the x direction. The radiometric dissimilarity between the two scenes is revealed by the low correlation values obtained for all bands. The two scenes are displayed in Fig. 4.3 with a zoom highlighting the measured mean displacement between the scenes. Figure 4.4 displays the scattergram calculated for each band of the two scenes within their domain of overlap and, contrary to the method used for calculating the regression coefficients, *without* suppressing regions contaminated by clouds. The presence of haze and clouds as well as illumination conditions and phenological differences (mid June to end of July in two different years) explain the rather low correlation values.

The generated ASCII files are stord in the directories `2006_COV12-QREG-XROI` and `2006_REF-QREG-XROI` where the letter `X` equals either `D` or `C` depending on the type of ROI used for the calculations. The files are organised using the level 18 of the European reference grid.

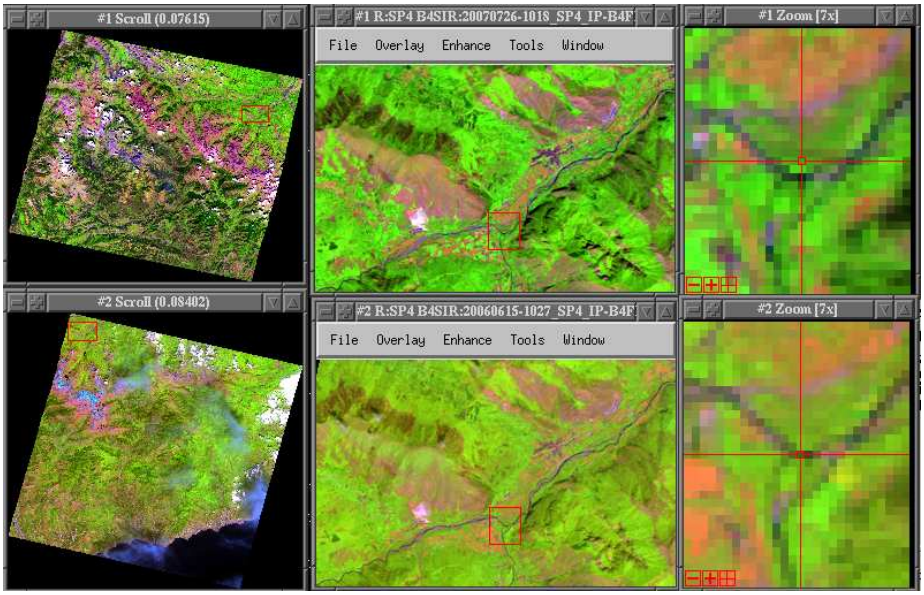


Figure 4.3: Example of two SPOT4 images in the French Alps with mean horizontal/vertical displacement equal to 17.5/-52m. The zoom section on the right highlights the vertical displacement. Measurements regarding the geometric and radiometric dissimilarity between these two scenes are given in Fig. 4.2.

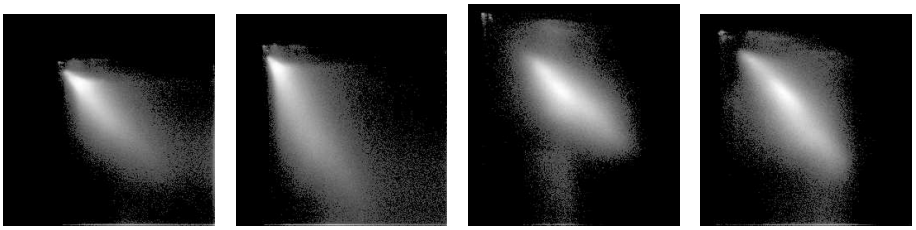


Figure 4.4: Scattergrams of bands 1 through 4 calculated within the domain of overlap of the image pair shown in Fig. 4.3. Measurements regarding the geometric and radiometric dissimilarity between these two scenes are given in Fig. 4.2.

4.5 Results

Geometric and radiometric consistency measurements described in the previous section were computed for all pairs of overlapping images using both DROIs and CROIs. For CROIs, the regions beyond the territory of the 38 participating countries plus a buffer of 5km (200 pixels at a resolution of 25m) have not been taken into account. A pair was deemed overlapping if at least 7 control points, in the sense of the threshold values detailed in Sec. 4.2, were found.

In this section, we present a series of synoptic tables digesting the information stored in the 55,566 files (29,447 for DROIs and 26,119 for CROIs). Wherever appropriate, these tables are given on a country and sensor basis. Beware that all country codes used in this chapter are referring to the codes used in the data delivered to JRC and therefore depart sometimes from normalised codes, see details in Chap 1 and appendix A.

4.5.1 Geometric consistency

An evalutation of the relative correspondence between overlapping pairs sorted on a sensor basis is summarised in tables 4.2–4.4 by indicating the mean, the RMSE, and the maximum of mean x-y displacements. These tables show that the geometric consistency, when considering pairs of identical sensors, increases in the following order: SP4, IL3, and SP5, the SP5 and IL3 values being very similar.

Measurements obtained for overlapping image pairs originating from the same country are summarised in table 4.5 for CROIs. This table indicates the average of the mean displacements, the average of the RMSE values as well as the maximum mean displacement among all selected overlapping pairs. It shows that the average of mean x-y displacements is below half a pixel for all countries. However, for almost all countries, there exists at least one pair of images showing a displacement of more than one pixel in either x or y directions. Also, one can observe that the geometric consistency of the within country pairs varies from one country to another. This variability is illustrated in Fig. 4.5 where the mean horizontal and vertical displacements of all pairs occurring in Iceland and Estonia are plotted.

Displacement measurements between an image originating from a country and all those originating from neighbouring country are presented for each country in appendix A.

4.5.2 Radiometric consistency

The frequency distribution (histogram) of the correlation values calculated for all overlapping pairs originating from the same sensor or between different sensors are shown in Fig. 4.6. A bin size of 0.05 has been used for producing these histograms. All intervals include the lower bound and not the upper bound (except the last one that also contains the upper bound). Note that there are many more pairs with a correlation value in the range $[0.95, 1.0]$ when considering pairs originating from the same sensor. This is mainly due to along path overlapping pairs acquired on the same day, i.e., originating from the same raw image.

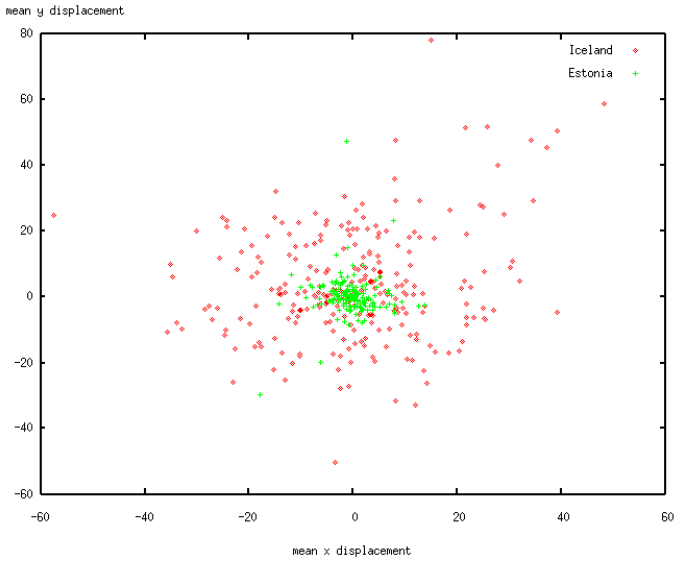


Figure 4.5: Mean displacements of all pairs of overlapping images occurring in Iceland (red diamonds) and Estonia (green crosses).

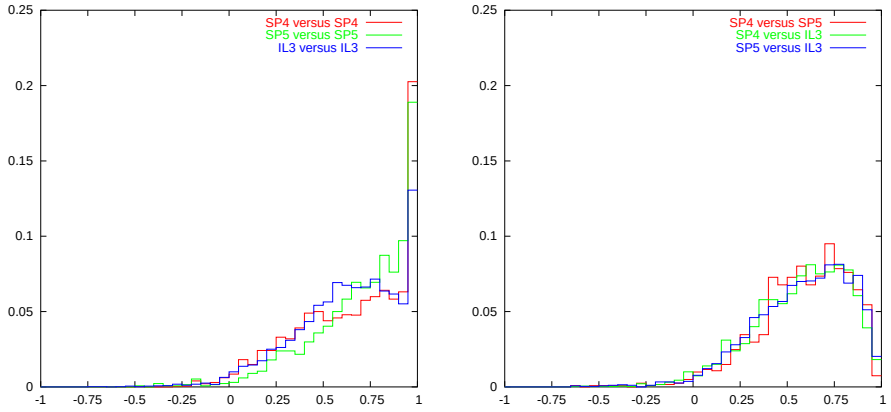


Figure 4.6: Histogram of the correlation values calculated for all overlapping pairs originating from the same sensor (left) or between different sensors (right). All calculations are based on DROIs.

Table 4.2: Between sensor mean of mean x-y displacements between all overlapping pairs based on DROIs and CROIs.

	DROI			CROI		
	SP4	SP5	IL3	SP4	SP5	IL3
SP4	8.6-7.0	8.4-6.9	6.6-7.4	6.5-7.2	6.9-7.0	6.5-6.7
SP5	6.6-7.4	6.5-7.2	4.4-5.0	4.3-4.8	5.0-5.8	4.6-5.5
IL3	6.9-7.0	6.5-6.7	5.0-5.8	4.6-5.5	4.6-5.4	4.3-5.1

Table 4.3: Between sensor mean of x-y RMSE displacements between all overlapping pairs based on DROIs and CROIs.

	DROI			CROI		
	SP4	SP5	IL3	SP4	SP5	IL3
SP4	11.5-10.0	11.3-9.9	10.2-10.9	10.1-10.7	10.7-10.7	10.3-10.4
SP5	10.2-10.9	10.1-10.7	6.3-7.3	6.2-7.2	8.3-9.3	7.9-9.1
IL3	10.7-10.7	10.3-10.4	8.3-9.3	7.9-9.1	8.1-8.9	7.7-8.5

Table 4.4: Between sensor maximum of the mean x-y displacements between all overlapping pairs based on DROIs and CROIs.

	DROI			CROI		
	SP4	SP5	IL3	SP4	SP5	IL3
SP4	111.8-78.3	111.8-82.8	61.4-58.8	61.4-58.7	79.8-70.2	79.8-80.2
SP5	61.4-58.8	61.4-58.7	52.5-50.5	52.5-49.0	67.8-85.0	53.4-85.0
IL3	79.8-70.2	79.8-80.2	67.8-85.0	53.4-85.0	70.8-67.3	67.7-65.4

Beyond the correlation values, the visual quality of the mosaic depends also on the coefficients of the linear regression computed for the top of atmosphere values. Ideally, the slope parameter should be close to 1 and the offset close to 0.

4.6 Conclusion

Measurements presented in this chapter enable to assess whether any given pair of overlapping images are in agreement in both geometry and radiometry. Therefore, when creating a mosaic, a quantitative evaluation of the degree of match between any pair of adjacent pieces of the mosaic will be straightforward as illustrated in Chap. 6. In particular, a visual representation of the geometric and radiometric consistency will be obtained by setting the boundary of the mosaic pieces to appropriate consistency measurements calculated between the corresponding overlapping image pairs.

Contrary to consistency measurements, accuracy measurements require the availability of a reference data set of known accuracy. Given the unavailability of a reference data set covering the entire territory of the 38 IMAGE-2006

Table 4.5: Within country geometric consistency assessment using CROIs.

CC	#pair$_i$	$\overline{\lvert X_i \rvert}$	$\overline{\lvert Y_i \rvert}$	$\overline{\text{RMSE-}X_i}$	$\overline{\text{RMSE-}Y_i}$	$\lvert \overline{X}_i \rvert_{\max}$	$\lvert \overline{Y}_i \rvert_{\max}$
al	95	10.5	5.8	12.8	8.1	49.4	24.7
at	375	4.4	6.7	7.0	9.4	22.4	45.7
ba	175	3.1	4.6	6.0	7.2	20.9	19.4
be	111	3.0	4.8	6.1	7.8	18.8	22.7
bg	464	7.6	5.3	11.4	8.9	44.5	26.9
ch	189	4.7	6.7	8.8	10.6	24.0	47.8
cs	338	7.3	5.6	10.8	9.4	35.7	30.7
cy	8	2.6	5.5	5.1	7.3	4.6	13.5
cz	263	4.8	3.8	7.4	6.8	20.2	26.0
de	1391	3.9	5.4	7.1	8.4	23.3	44.9
dk	308	4.8	3.7	10.2	9.9	25.6	29.0
ee	156	3.1	3.0	6.4	6.1	17.9	29.8
es	1083	3.7	4.7	6.2	7.4	22.3	37.6
fi	1042	10.4	9.4	13.0	11.9	111.8	58.9
fr	1696	5.1	7.5	8.2	10.6	28.3	53.8
gb	1704	5.5	5.0	10.0	10.4	33.4	30.7
gr	814	6.2	4.9	8.8	7.4	38.5	40.4
hr	158	4.0	5.6	7.6	8.8	19.1	53.2
hu	281	7.0	5.4	11.3	10.4	26.8	31.2
ie	464	5.1	6.9	9.3	11.3	24.5	41.9
is	241	11.3	12.7	15.4	16.8	48.2	78.3
it	1179	6.0	6.7	9.0	9.8	54.1	85.0
li	1	0.6	9.5	3.4	10.6	0.6	9.5
lt	290	5.6	5.6	8.4	8.6	22.6	27.5
lu	13	2.9	4.3	5.2	6.5	6.4	12.2
lv	161	4.4	3.5	6.7	5.8	21.5	14.2
mc	57	8.0	6.0	10.9	9.3	29.8	24.6
me	35	6.0	4.4	8.4	7.5	24.5	15.0
mt	1	9.7	4.5	16.5	25.2	9.7	4.5
ni	81	5.8	6.0	10.7	11.7	23.3	26.4
nl	295	5.0	4.7	11.3	11.0	31.5	31.2
no	1733	6.0	8.4	10.3	11.9	61.4	51.9
pl	841	5.0	3.4	7.9	6.5	26.4	20.2
pt	242	5.1	4.2	7.3	6.5	28.5	36.0
ro	824	7.3	5.2	11.8	10.3	43.7	26.6
se	1304	6.0	7.3	8.1	9.4	43.9	52.2
si	78	4.4	4.7	6.9	7.1	19.6	19.7
sk	138	5.6	3.4	9.3	7.4	27.0	17.7
tr	1828	4.3	4.2	7.1	7.0	34.6	23.9

participating countries, accuracy measurements are beyond the scope of this chapter. Nevertheless, one should keep in mind that inconsistencies revealed by a consistency check always indicate the presence of inaccurate data. For example if the norm of the displacement vector between 2 images is equal to 2 pixels, the minimum of the norms of the two displacement vectors calculated for each image with respect to ground truth is equal to 1 pixel. This 'best' case (given a measured disagreement of 2 pixels between the 2 images) occurs in the unlikely event where each image would be translated by 1 pixel, but in opposite directions with respect to ground truth.

Chapter 5

Mosaicing methodology

Pierre Soille and Conrad Bielski

Spatial Data Infrastructure Unit
Institute for Environment and Sustainability
DG Joint Research Centre, European Commission

5.1 Introduction

To construct an image of larger field of view than what could be obtained with a single image acquisition, two or more image acquisitions need to be combined. The resulting image is called a mosaic in the sense that it is made of fragments of different images. For example, with IRS or SPOT imagery, a country such as Belgium cannot be covered by a single image. In addition, if one desires cloud cover not to exceed a given percentage, some regions may need to be covered more than once. Figure 5.1 shows that 15 IRS and SPOT scenes were necessary in the first coverage of IMAGE-2006 to cover the territory of Belgium while securing that cloud cover does not exceed 5%. In this figure, the images have been loaded in an arbitrary order so that the value of any given pixel corresponds to the value given by the last image providing values for this pixel. That is, one would obtain the same result by superimposing printed copies of the images using the same order as the one used for loading the images. Because the image have been orthorectified and projected beforehand, data values occupy only a portion of the coloured rectangle frame of each image (small frames for SPOT, large frames for IRS images).

The mosaicing of a set of spatially overlapping images comes down to determine a unique value for each pixel that is covered by more than one image. Beyond trivial (and arbitrary) methods such as the one used for creating Fig. 5.1, this can be achieved by pixel or region based methods. With pixel based methods, the output value of a pixel is determined with the sole knowledge of the input values of this pixel. By contrast, with region based methods, the output value of a pixel depends on some spatial relationships with other pixels.

Among pixel based methods, those relying on a selection procedure are often

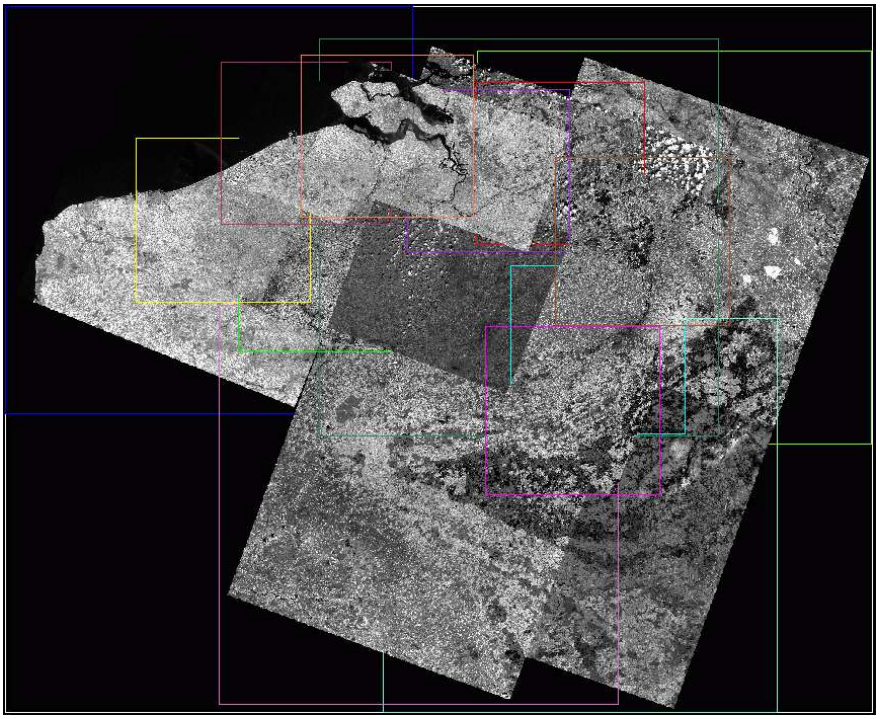


Figure 5.1: The 14 scenes delivered for Belgium composed according to the (arbitrary) order in which they have been loaded. The frames indicate the relative positions of these overlapping scenes. The displayed data corresponds to raw digital numbers of band 4.

performing better than those relying on linear combinations of the available values. A simple per pixel selection method is to define as output value of a pixel the minimum or maximum value of the available input values. This approach is of limited interest for creating a seamless mosaic because any variation of reflectance values due to atmospherical or seasonal effects leads to blocky structures in the output mosaic. Still, the point-wise minimum and maximum composition rules are of practical interest. For instance, the minimum composition rule generates a mosaic showing the least possible amount of cloud cover because clouds have usually higher reflectance values than most land cover classes in the optical domain. Also, because atmospheric effects such as haze and thin clouds increase the top-of-atmosphere (TOA) reflectance values, the minimum mosaic tends to select scenes that are less corrupted by atmospheric effects. Similarly, shadows and water bodies as well as all objects having low reflectance values are highlighted in the minimum mosaic. The maximum mosaic, on the other hand, reveals all clouds occurring in all overlapping images and tend to give precedence to TOA reflectance values most affected by atmospheric conditions. Nevertheless, the minimum and maximum mosaics are useful for visual control purposes. For instance, any cloudy area appearing in the minimum mosaic can be considered as an area permanently covered by clouds given the available imagery, i.e., it will appear in any other mosaic based on the same set of input images. The maximum mosaic is also useful for visually assessing the perfor-

mance of a cloud detection algorithm: the union of all cloud masks extracted for each individual image should match all clouds visible in the maximum mosaic. This principle was used for creating the cloud galleries presented in Sec. 3.4, page 58). All effects described above can be observed in the point-wise minimum and maximum mosaics displayed in Fig. 5.2. In this figure, the mosaics have been rendered using the Top of Atmosphere (TOA) reflectance values of the 4th, 3rd, and 1st bands for, respectively, the red, green, and blue channels.

Region based methods rely on a partition of the domains of overlap. Each segment of the partition is associated with an index value indicating from which image the values of the mosaic must be selected from within this segment. Hence, each segment with its associated index value can be considered as a decision region. Because decisions are taken at the level of a region, region based methods lead to mosaics showing a higher spatial coherence than pixel based methods.

The IMAGE-2006 mosaics were produced thanks to a region based method called morphological image compositing [34]. This method partitions automatically the domains of overlap. It is best introduced in the simplest case consisting of two partly overlapping images that need to be composed to form a unique image. In this case, the problem comes down to partitioning their domain of overlap into two regions indicating which image should be selected within these regions. The interface between these two regions define a cut or seam line because on one side of this line an image is used while the other one is used on the other side. Therefore, the seam line should be placed so as to diminish as much as possible the visual detection of the transition from one image to another in the resulting mosaic. That is, cuts revealing noticeable mismatch in tones and position should be avoided. In the early days of image mosaicing, this was achieved by making a razor cut following terrain structures such as roads, railroads, edges of fields, or other lines of definite tone demarcation [45, p. 242]. This old-fashioned method is illustrated in Fig. 5.3 and was still used in the first continental mosaic, namely the first Landsat mosaic of North America in 1976 [14] consisting of 569 Landsat images pieced together by technicians during 4 months. Morphological image compositing [34] mimics human operators by automatically defining seam lines following the most salient object boundaries occurring within the domain of overlap. This is achieved thanks to a region growing process [48]. The growth is initiated by defining two seeds corresponding to the regions where there is no overlap (one seed for each image). The actual growth proceeds within the domain of overlap until it is completely covered. The domains reached by each seed indicate which image should be used within these domains. To ensure that the boundary of these domains matches salient image structures, the growth is constrained by an image whose intensity values are proportional to variations in tone within a neighbourhood centred at each pixel of the overlapping domain. That is, when variations are high, the speed of the growth process is decreasing and vice versa. The propagation is computed in such a way that growing regions actually meet along lines of high tonal variations corresponding to the most salient image structures. Morphological image compositing can be extended to an arbitrary number of images while minimising the occurrence of undesirable objects such as clouds.

The goal of this chapter is to present a detailed description of the methods developed to this aim. It is organised as follows. Section 5.2 presents a formal description of morphological compositing for the direct (in place) processing of

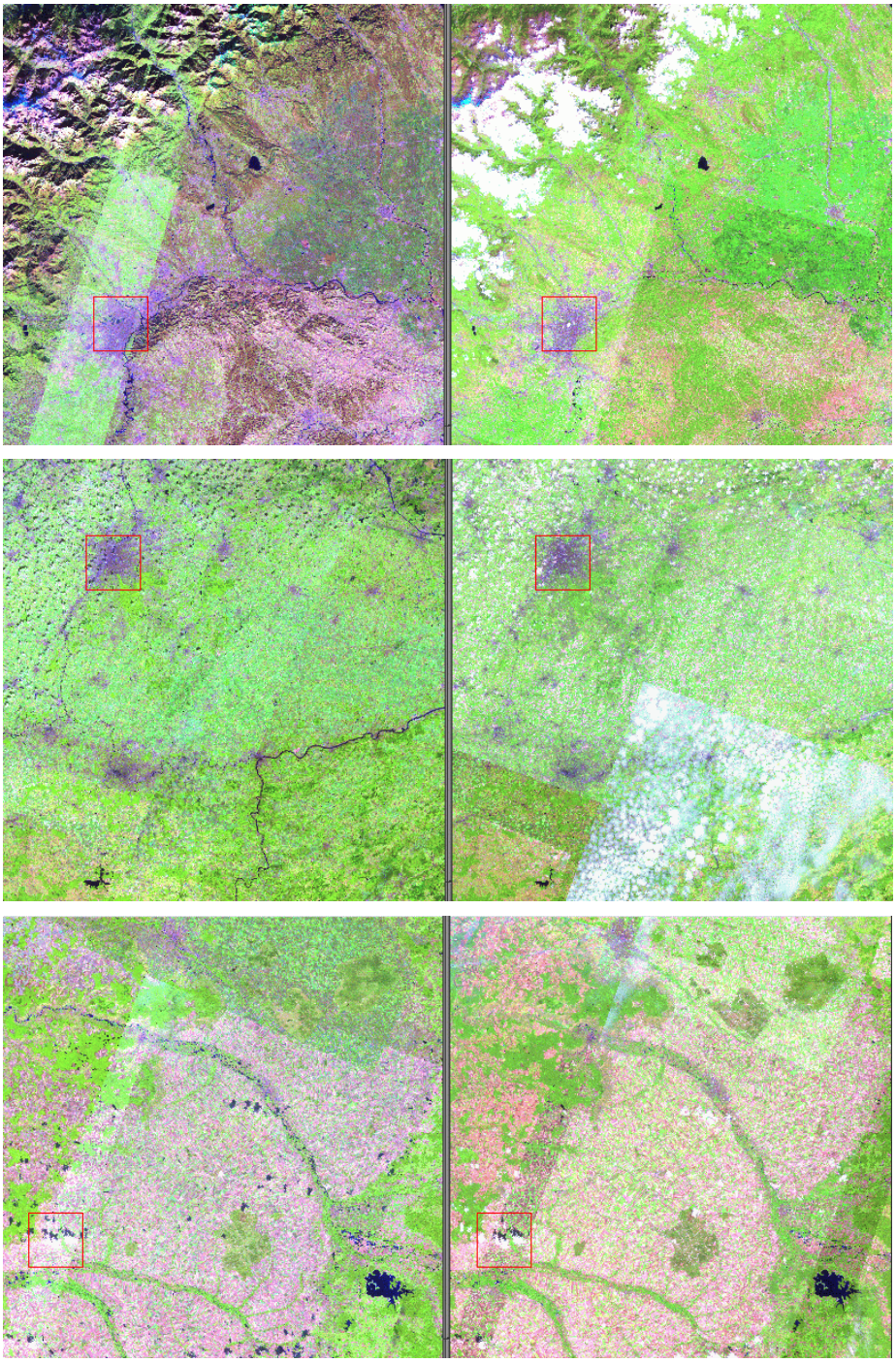


Figure 5.2: Comparison between the point-wise minimum (left) and the point-wise maximum (right) mosaics computed on the basis of all imagery of the first coverage of IMAGE-2006. The areas framed by a red box represent Turin (top), Brussels (middle), and a permanent cloud South of the Seine river (bottom).

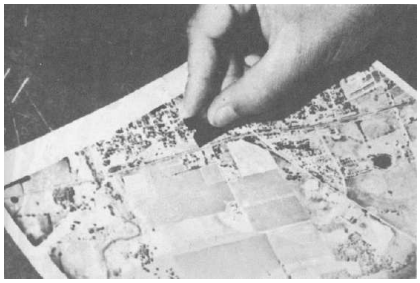


Figure 5.3: Making the razor cut or seam line by following lines of definite tone demarcation. Source: Wolf [45, p. 241].

the images to compose. The direct processing of a large number of images is rapidly limited by the available computer random access memory. Section 5.3 shows that this limitation can be addressed thanks to sequential processing. For a very large data sets such as the IMAGE-2006 imagery, sequential processing would require too much time. Section 5.4 addresses this issue by introducing a parallel algorithm enabling the distribution of the mosaicing task to a series of computers. Concluding remarks are presented in Sec. 5.5.

5.2 Direct processing

Mosaicing consists in merging a sequence of spatially overlapping images so as to create a unique image whose spatial extent is equal to the smallest rectangle enclosing all input images. Direct processing assumes that enough memory is available to perform all computations on temporary images whose extent is equal to that of the output mosaic. The fundamental principles of morphological image compositing were first introduced in the case of direct processing. They are recalled hereafter, first for the simplest case of 2 images to compose and then for an arbitrary number of images.

5.2.1 Composing 2 images

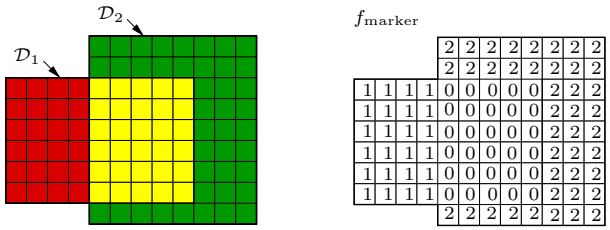
Let $\mathcal{D}_1$ and $\mathcal{D}_2$ denote the definition domain of two digital images f_1 and f_2. Let assume that these images partly overlap, i.e., $\mathcal{D}_1 \cap \mathcal{D}_2 \neq \emptyset$, and $\mathcal{D}_2 \neq \mathcal{D}_1$. The goal is to create an image f whose definition domain $\mathcal{D}_f$ equals $\mathcal{D}_1 \cup \mathcal{D}_2$. Note that, in practice, this image is represented in computer memory by an image whose extent is equal to the domain of the smallest rectangle enclosing $\mathcal{D}_1 \cup \mathcal{D}_2$. The morphological composition method is based on a region growing process starting from markers and driven by a mask image.

Marker image

The marker image whose definition domain is equal to $\mathcal{D}_f$ is defined as follows for each pixel $\mathbf{x}$:

$$\Big[f_{\text{marker}}\Big](\mathbf{x}) = \begin{cases} 1, & \text{if } \mathbf{x} \in \mathcal{D}_1 \text{ and } \mathbf{x} \notin \mathcal{D}_2, \\ 2, & \text{if } \mathbf{x} \in \mathcal{D}_2 \text{ and } \mathbf{x} \notin \mathcal{D}_1, \\ 0, & \text{otherwise (i.e., no marker).} \end{cases}$$

The definition domains of two overlapping images and the corresponding marker image are illustrated in Fig. 5.4. That is, when creating the mosaic, for the pixels



				2	2	2	2	2	2
				2	2	2	2	2	2
1	1	1	1	0	0	0	0	2	2
1	1	1	1	0	0	0	0	2	2
1	1	1	1	0	0	0	0	2	2
1	1	1	1	0	0	0	0	2	2
1	1	1	1	0	0	0	0	2	2
1	1	1	1	0	0	0	0	2	2
				2	2	2	2	2	2

Figure 5.4: Left: definition domains $\mathcal{D}_1$ and $\mathcal{D}_2$ of two overlapping images. The yellow region defines the domain of overlap (combination of the fundamental colours used for each individual definition domain). Right: the corresponding marker image f_{marker} with its numerical values.

Figure 5.5: Two overlapping images and the corresponding marker image. Left: excerpt from `20060818-1041 IL3 FR 3963524448-BB`. Middle: excerpt from `20050715-1041 SP5 FR 3931424292-DB`. Right: corresponding marker image.

with value 1 (resp. 2) of the marker image, there is no choice but select the value of the image f_1 (resp. f_2) since only one value is available for these pixels. The problems reduces to determine from which image the values of the 0-valued pixels should originate. Indeed, two input values are available for these pixels. Two overlapping images and their corresponding marker image are illustrated in Fig. 5.5. Note the discrepancies in land cover use in the overlapping domain (temporal difference of 13 months between the two images). It follows that any mosaicing whose seam lines would not follow the boundaries of objects visible in *both* images would introduce undesirable discontinuities.

Mask image

The mask image is used to direct the growth of the regions where there is no overlap (i.e., regions defined with non-zero values in the marker image described in the previous section). By definition, the growth of the markers must be restricted to the domain of overlap so that the mask image needs to be defined solely within this domain. If it would be simply defined as an image with constant values, the growth would be isotropic and lead to regions corresponding to the (geodesic) influence zones [33] of the markers within the domain of overlap. That is, in the example of Fig. 5.5, the red marker would grow downwards and the green one upwards so that they would fill the rectangular overlapping domain (yellow) and meet exactly along its horizontal axis. Obviously, this would generate a visible seam line in the corresponding mosaic because it would not follow the boundaries of actual image objects. This can be solved by defining

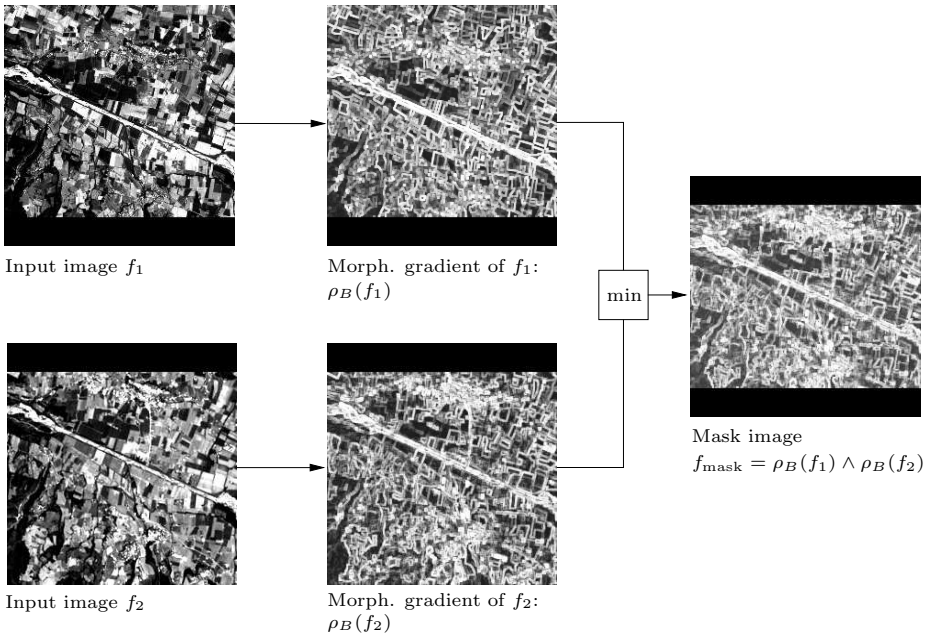


Figure 5.6: Generation of the mask image used for directing the growth of the markers.

a mask image highlighting the boundaries of the image objects. Because object boundaries correspond to regions of high tonal variations, they can be enhanced by computing the difference between the highest and lowest grey level values within an elementary neighbourhood centred on each pixel. This definition is known as the morphological gradient ρ within a neighbourhood B [32]. To ensure that only tonal variations occur in *both* images at the same locations, the morphological gradient of both images are computed in parallel and then combined through the point-wise minimum operation $\wedge$:

$$f_{\text{mask}} = \rho_B(f_1) \wedge \rho_B(f_2).$$

This is illustrated in Fig. 5.6. When computing the gradient of each individual image, pixels that are outside the data ROI of the image must be ignored to avoid spurious high gradient values along the boundaries of the data ROIs.

From marker and mask images to decision regions

Once the marker and mask images have been defined, the overlapping domain needs to be partitioned into two decision regions, one for each marker. This can be achieved by a morphological transformation known as the watersheds from markers [38, 43]. That is, the watersheds of the mask image are computed using the markers as attractors. The resulting catchment basins partition the domain of overlap into two influence zones (decision regions) indicating which image should be used in each zone. The resulting puzzle pieces are then combined to create the desired mosaic. All these steps are illustrated in Fig. 5.7.

In mathematical terms, composition of the input images is achieved with the

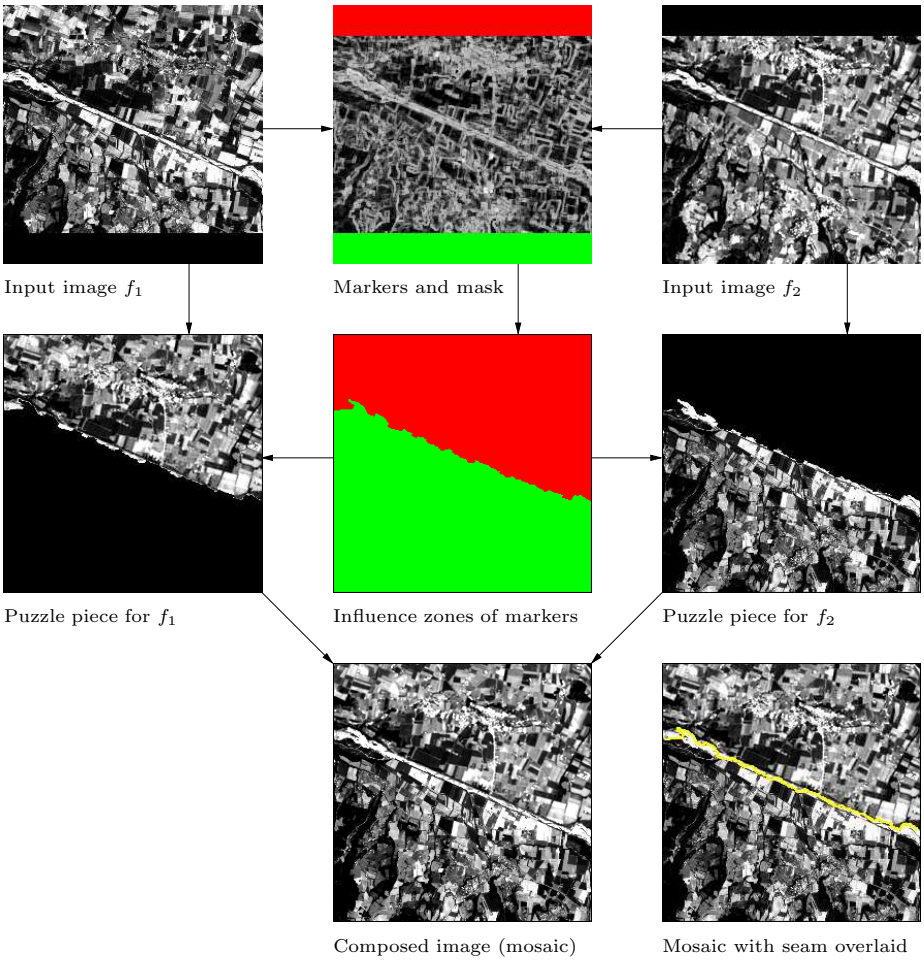


Figure 5.7: Morphological image compositing of two overlapping images: flow diagram (see text for details).

following decision rule:

$$f(\mathbf{x}) = f_i(\mathbf{x}),$$

where the value of the subscript i is defined by the label value of the influence zone IZ at position $\mathbf{x}$:

$$i = \Big[\mathrm{IZ}_{f_{\mathrm{marker}}}(f_{\mathrm{mask}})\Big](\mathbf{x}).$$

Removing specific objects

The methodology can be easily extended to remove specific object such as clouds. The first step is to detected in each image all objects that are undesirable. For instance, cloud detection on SPOT and IRS imagery is detailed in Chap. 3. The marker image is then updated to take into account the binary masks of the undesirable objects. For example, if a cloud is detected in an image, its extent defines an additional marker for the second image (from which the region growing procedure is also initiated). By doing so, provided that no cloud appears at the same location in the second image, one makes sure that the cloud will not appear in the mosaic.

Formally, denoting by M_i the binary mask of undesirable objects occurring in the image f_i, the marker image is defined as follows:

$$\Big[f_{\mathrm{marker}}\Big](\mathbf{x}) = \begin{cases} 1, & \text{if } \mathbf{x} \in (\mathcal{D}_1 \cap \overline{\mathcal{D}_2}) \text{ or } \Big[\mathbf{x} \in (\mathcal{D}_1 \cap \mathcal{D}_2) \text{ and } \mathbf{x} \in (M_2 \cap \overline{M_1})\Big], \\ 2, & \text{if } \mathbf{x} \in (\mathcal{D}_2 \cap \overline{\mathcal{D}_1}) \text{ or } \Big[\mathbf{x} \in (\mathcal{D}_1 \cap \mathcal{D}_2) \text{ and } \mathbf{x} \in (M_1 \cap \overline{M_2})\Big], \\ 0, & \text{otherwise (i.e., no marker),} \end{cases}$$

where the line over a set refers to the complement of this set. That is, in the case of satellite images, clouds (or any other undesirable object) occurring in an image are suppressed by considering markers of the other image at their positions (provided that clouds do not occur at the same position in the other image). The watershed-based propagation of the markers ensures that the resulting seam lines are following the boundaries of relevant image objects present in both images.

5.2.2 From two to an arbitrary number of images

The procedure described in Sec. 5.2.1 for the simplest case of 2 images can be extended to the mosaicing of an arbitrary image number n. This extension is described in the following subsections.

The overlap level image g

The key to the extension from 2 to n images consists in processing the domain of overlaps by increasing overlapping level. The overlap level of a given pixel indicates the number of images covering this pixel.

Let n denote the number of spatially overlapping images. Let $\mathcal{D}_f$ denote the definition domain of the composed image: $\mathcal{D}_f = \cup_i \mathcal{D}_i$. The overlap level image g indicates how many images are overlapping any given pixel $\mathbf{x}$ of $\mathcal{D}_f$:

$$g(\mathbf{x}) = \mathrm{card}\{i \mid \mathbf{x} \in \mathcal{D}_i\}. \tag{5.1}$$

For example, the left diagram of Fig. 5.8 shows the definition domains of 3 overlapping images. Fundamental colours and their combinations have been

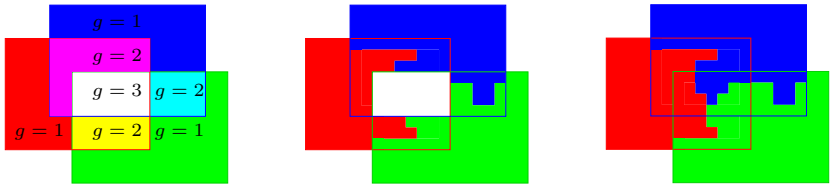


Figure 5.8: Three overlapping images. Left: overlap level image g. The yellow, magenta, cyan, and white regions define the domain of overlap (combination of the colours used for each individual definition domain where no o overlap occurs). Middle: after region growing in the domains where two images overlap. Right: after subsequent region growing in the domain where all 3 image overlap. Adapted from [34, Fig. 3].

used to highlight the domains of overlap. The g image is equal to one for the fundamental colours, 2 for the 3 compositions of 2 fundamental colours, and 3 for the central white domain where all 3 images overlap.

Mask image

The mask image is defined as the point-wise minimum between the morphological gradients of all input images:

$$f_{\text{mask}} = \bigwedge_i \rho_B(f_i). \tag{5.2}$$

By doing so, cut lines will naturally follow the the boundaries of object that are visible in *all* available images. When computing the gradient of each individual image, pixels that are outside the data ROI of the image must be ignored to avoid spurious high gradient values along the boundaries of the data ROIs.

Iterative procedure

When composing an arbitrary number of images, markers grow through an iterative procedure from an initial state (matching the regions where there is no overlap) to a final state where they partition the domains of overlap into a series of decision regions.

The initial marker image is analogous to that introduced for the composition of an image pair. The markers correspond to the regions where there is no overlap (red, green, and blue region in the left image of Fig. 5.8). They are then propagated within the regions of overlap level 2 (see middle image of Fig. 5.8), and then to the subsequent levels until the maximum degree of overlap is reached. In the example of Fig. 5.8, this occurs at level 3 (see right image). At each step, the propagation is directed by the mask image defined in the previous section.

This iterative procedure is illustrated in Fig. 5.8 and can be formalised as follows. Let k refers to an arbitrary iteration level ($1 \leq k < g_{\max}$). The k-th iteration processes the pixels that have on overlap level equal to k. Accordingly, we denote by $f^{(k)}$ the values of f that are defined at the end of the k-th iteration. Consequently, the definition domain of $f^{(k)}$ is equal to those pixels of the overlap

level image whose values are less than or equal to k:

$$\mathcal{D}_{f^{(k)}} = \{\mathbf{x} \mid g(\mathbf{x}) \leq k\}.$$

Let $\mathcal{D}_i^{(k)}$ refer to those pixels of $\mathcal{D}_{f^{(k)}}$ whose values originate from the input image f_i:

$$\mathcal{D}_i^{(k)} = \{\mathbf{x} \in \mathcal{D}_{f^{(k)}} \mid f^{(k)}(\mathbf{x}) \leftarrow f_i(\mathbf{x})\}.$$

Initially, $k = 1$ and therefore $\mathcal{D}_i^{(1)}$ is trivially defined since there is only one image covering each pixel $\mathbf{x}$. For actual overlaps (i.e., $k > 1$, the definition of $\mathcal{D}_i^{(k)}$ is achieved thanks to the computation of the decision regions by the marker-controlled segmentation described in Sec. 5.2.1. The marker image at iteration $k > 1$ is defined by the decision regions already determined during the previous iteration(s):

$$f_{\text{marker}}^{(k)}(\mathbf{x}) = \begin{cases} i, & \text{if } \mathbf{x} \in \mathcal{D}_i^{(k-1)}, \\ 0, & \text{otherwise (i.e., no marker).} \end{cases}$$

The decision rule is then be formulated as follows:

$$f^{(k)}(\mathbf{x}) = f_i(\mathbf{x}) \text{ where } i = \left[\text{IZ}_{f_{\text{marker}}^{(k)}}(f_{\text{mask}})\right](\mathbf{x}),$$

the propagation if the markers being restricted to $\mathcal{D}_{f^{(k)}}$. Note that during the computation of the influence zones, it is necessary to check that markers only propagates within the definition domain of the image they correspond to. The iterative procedure stops when k reaches the largest overlap level $g_{\max}$.

Removal of specific objects

The procedure for the removing specific objects in the case of two overlapping images (see Sec. 5.2.1) is more complex when a higher number of images overlap. Indeed, assuming that a given pixel belongs to an object detected solely in one image and that should not appear in the composed image, there is no direct selection rule to choose among the remaining $n-1$ pixels values. This problem can be solved by remapping the original image indices $i \in \{1, \ldots, n\}$ to 2^{i-1} so as to produce unique values when summing two of more indices (the sum is then equivalent to a bit-wise OR operation). If an undesired object such as a cloud-shadow complex occurs in a region of overlap, one needs to create a marker whose spatial extent corresponds to this object and whose value equals the bit-wise OR operation between the remapped indices of all images overlapping this object, except the one containing this cloud-shadow complex. The resulting index is referred to as *composite* index because it represents the union of two or more image indices. For instance, in the example depicted in Fig. 5.8, in case a cloud-shadow complex of the second image would occur in the region where all three images overlap, a marker with a composite index value equal to 5 would be generated where this complex occurs. Indeed, the sum of the powers of 2 of the index values corresponding to the first and third images equals $2^{1-1} + 2^{3-1} = 5$. In the sequel, the set of indices used for generating a composite index is called its generator set. Indices referring to a unique image are called *plain* indices.

By definition, plain indices are always powers of 2. A formal definition of the marker image follows:

$$f^{(k)}_{\text{marker}}(\mathbf{x}) = \begin{cases} c, & \text{if } \mathbf{x} \in \mathcal{D}_c^{(k-1)}, \\ \sum_i \{2^i \mid \mathbf{x} \in \mathcal{D}_i \text{ and } x \notin M_i\}, & \text{if } g(\mathbf{x}) = k \text{ and } \exists\, j \mid \mathbf{x} \in M_j, \\ 0, & \text{otherwise (i.e., no marker).} \end{cases}$$

$c \in \{1, \ldots, 2^n - 1\}$ and $\mathcal{D}_c^{(k-1)} = \left\{ \mathbf{x} \in \mathcal{D}_f^{(k-1)} \;\middle|\; \left[\text{IZ}_{f^{(k-1)}_{\text{marker}}}(f_{\text{mask}})\right](\mathbf{x}) = c \right\}$. At each iteration, the composite label values of the influence zones originating from specific structures to remove must be mapped to plain label values. This is achieved by using a majority rule: each influence zone with a composite label value is scanned and is set to the plain label value of its adjacent influence zone with which it share the largest number of boundary pixels. If such a region is not found (e.g., all adjacent regions are labelled with composite labels) then the smallest plain label forming the composite label is selected.

5.3 Sequential processing

Sequential processing enables the processing of one image at a time without the need to hold in random across memory an image whose size is equal to the smallest rectangle encompassing the domains of all input images. Still, it is desirable to produce results that are independent of the processing order. Indeed, order independence ensures that the same results are obtained whatever the order in which the images are processed so that results are exactly reproducible. The pursuit of this goal leads to the idea of sequential order independent image compositing. This idea was first briefly presented in [13] and is summarised hereafter.

5.3.1 The concept of overlap matrix

When processing a given image, it is convenient to know at any time the set of images whose definition domains overlap this image. This can be achieved by computing once for all the so-called *overlap matrix*. This is a symmetrical $n \times n$ matrix indicating whether the definition domains of an arbitrary image pair overlap or not:

$$m_{i,j} = \begin{cases} 1, & \text{if } \mathcal{D}_i \cap \mathcal{D}_j \neq \emptyset, \\ 0, & \text{otherwise.} \end{cases} \tag{5.3}$$

In general, owing to the arbitrary shape of the definition domains $\mathcal{D}_i$, the overlap matrix cannot be constructed on the sole knowledge of the frame size and positioning of the upper left corner in the reference coordinate system. Indeed, the footprint of a georeferenced satellite image does not cover the full image frame (see example in Fig 5.1). The domain (footprint) covered by data values is called data region of interest (data ROI). It has usually a trapezoidal shape or more complex shape in case the data values values are deemed reliable only in sub-domain of the full data ROI domain. Hence, for all pairs of images whose frames intersect, we need to compute the intersection between their actual definition domains to assess whether these intersect or not.

5.3.2 Ordered propagation

Beyond the information available through the overlap matrix, the number of overlapping images (called *overlap level* hereafter) for any given pixel must be known. Indeed, seam lines are detected in an ordered fashion, starting with all regions whose overlap level is equal to 2 and proceeding to the subsequent level until the maximum number of overlap is reached.

We then proceed as follows. The $n \times n$ overlap matrix is scanned row by row. The index of the row defines the current *anchor image*. Actual processing is restricted to the region defined by its frame: we assume that the actual definition domain (i.e., data ROI) is buffered by the image frame). Within this region, the least overlap level greater than 1 defines the *current overlap level* denoted by k. We also track whether an overlap level higher than the current level occurs. This is used to determine whether the overlap matrix needs to be scanned again later on. The morphological compositing routine [34] is then called while restricting its effect to the processing of those regions whose overlap level is equal to k. The routine assigns each pixel of these regions to a unique source image so that the definition domain of the anchor image and those of the images intersecting it can be updated accordingly. This update can only remove some parts of these input definition domains since it concerns regions where more than one image was competing for the same domain. This procedure secures order independence in the sense that identical results are obtained whatever order the anchor images are processed.

5.4 Parallel processing

The sequential algorithm described in the previous section allows for the processing of an arbitrary number of scenes. However, because these scenes must be processed one after the other, this is not a solution for processing very large image data sets ($n > 100$) such as continental or global coverages. In this section, we propose an order independent parallel algorithm allowing for well chosen individual images to be processed in parallel. The concepts of overlapping matrix and ordered processing by increasing level of overlap level originally developed for sequential processing are also used for parallel processing. Section 5.4.1 introduces the notion of independent sets of anchor images. The parallel algorithm as such is described in Sec. 5.4.2.

5.4.1 Independent sets of anchor images

The key to parallel processing is to recognise that two scenes can be processed in parallel if and only if they are independent, i.e., if the processing of the first one at the current processing level is independent of the outcome of the processing of the second one at the same processing level. This happens if there is an empty intersection between (i) the definition domain (ROI) of the first scene unioned with the definition domains (ROIs) of all its overlapping scenes and (ii) the definition domain of the second scene unioned with the definition domains (ROIs) of all its overlapping scenes. A list of independent images is such that any pair of images of the list are independent. A largest list of independent images is any list of independent images that cannot be extended without violating the independence condition. The search for *the* largest list(s)

of independent images that can be extracted from the image list may prove very difficult (a brute-force search would require testing all permutations of the input image list). Therefore, in practice, it is acceptable to generate *a* largest list. Such a list can be obtained by scanning the image list and insert the current image in the independent list if it is independent of all images already inserted in the independent list. The resulting list of independent images is not sufficient because it does not contain all input images. Therefore, the procedure must be repeated so as to create additional lists of independent images until all images are selected. That is, as soon as an image is selected in a list, it is flagged as not selectable during the creation of subsequent lists. In addition, because images must be processed in increasing order of overlap level, the creation of lists of independent images must be performed for each successive processing level. To speed up the creation of the lists, when creating the overlapping matrix, the maximum overlap level of each image is computed. This prevents from inserting any image whose maximum overlap level exceeds the current overlap level.

5.4.2 Ordered processing

Ordered processing is achieved by processing images in parallel for increasing overlap levels. Initially, all regions where only 2 images overlap are processed. This step is itself achieved by processing in parallel a largest list of independent images and then a subsequent list of independent images until all images have been processed. The algorithm proceeds by processing all regions where 3 and only 3 images overlap, and so forth until the maximum level of overlap $g_{\max}$ is reached.

This algorithm can be run on a batch queueing system. This type of systems distributes automatically the composition of the independent images held in the list plist to the available processors. Once all images of the independent list have been processed, the subsequent list is sent to the queueing system and so forth until all images have been processed for all overlap levels. Figure 5.9 shows a grid engine composed of 16 blades with two CPUs each. This system allows for the definition of cut lines for a pan-European coverage at 25m resolution (more than 1,500 images) in less than 12 hours. This modern razor can be compared with the one used in the seventies (see Fig. 5.3).

5.5 Conclusion

The morphological image compositing proposed in [34] allows for the automatic definition of seam lines during mosaicing. In this chapter two algorithms enabling the processing of large data sets have been put forward. The first permits sequential processing so that it is not necessary to hold in random access memory temporary images whose size are equal to that of the target mosaic. The second enables the processing of several images in parallel. This latter algorithm is an asset for mosaicing large data sets such as the IMAGE-2006 imagery covering the entire territory of the European Union plus 11 additional countries. The resulting cloud free mosaics are described in Chap. 6.

Figure 5.9: Making the razor cut with a computer farm consisting of 16 blades containing each 2 CPUs and allowing for the definition of cuts lines of more than 1,500 images in less than 12 hours. Compare with Fig. 5.3.

Chapter 6

Mosaicing results

Pierre Soille and Conrad Bielski

Spatial Data Infrastructure Unit
Institute for Environment and Sustainability
DG Joint Research Centre, European Commission

6.1 Introduction

The mosaicing methodology described in Chap. 5 was applied to the IMAGE-2006 data sets to generate a series of mosaics. Each mosaic is supplemented with a range of additional information layers of interest to subsequent information extraction. Actually, the core information layer does not contain the multispectral imagery but simply the labelled decision regions together with the correspondence table indicating which image should be selected for each label value of the decision regions. Hence, the determination of the origin of the pixel values of any pixel of the mosaiced multispectral imagery is straightforward. Furthermore, the decision regions can be directly mapped, thanks to appropriate look-up-tables, to sensor type, year of acquisition, or any other attribute varying from one image to another.

This chapter is organised as follows. A description of each type of each stored information layer is given in Sec. 6.2. The mosaics of coverages 1 and 2 are described in Secs. 6.3 and 6.4 respectively. A brief conclusion is presented in Sec. 6.5.

6.2 Information layers

The produced mosaics are organised in directories indicating the coverage name and type of ROI used (either DROI or CROI). Images for a given coverage were retrieved from the reference coverage by searching for all header files containing this coverage in their `data_Set` field, see sample header file in Fig. 1.5 of Chap. 1. Indeed, contrary to image searches conducted directly on coverage 1 or 2, image

searches conducted on the reference coverage always lead a set of unique images because multiple deliveries of the same scene were merged when creating the reference coverage, see Sec. 1.6. To emphasise that images were retrieved from the reference coverage, the string `REF` is included in the directory name of the mosaic products. For example, `2006_REF-MOS-COV1-CROI` corresponds to the mosaic product of all imagery of the first coverage retrieved from the reference coverage and calculated on the basis of CROIs.

Each mosaic is then stored in directories following the level 18 of the European Reference Grid [2]. Each level 18 (square) tile covers a region of 250,000 km^2 stored in a raster file of $20{,}000 \times 20{,}000$ pixels at a spatial resolution of 25 m. For instance, Switzerland falls fully in the level 18 tile named `42-C`. For convenience, the chosen convention for naming level 18 information layers departs somewhat from the specifications put forward in [40]. The `date-time` field has been set to the median of the dates of the whole input imagery so that the same value is used for all tiles[1]. The centre pixel field has been substituted by the level 18 tile name itself, e.g., `42-C`. The `type` sub-field of the `type-identifier` field is set to `MXT` to emphasise that the tiles consist of a mixture of more than one image.

Although all stored layers appear together in a given level 18 directory, they can be naturally grouped into three categories: (i) base layers, (ii) mosaic layers, and (iii) quality layers. Each category is described in a separate section hereafter.

6.2.1 Base layers

Base layers are information layers that were generated independently of the mosaicing results using point-wise operations between the input imagery, their ROIs, or their cloud masks. They consist of 4 layers:

1–2. point-wise minimum (`date_MXT_CM-INFn_tile.tif`) and point-wise maximum (`date_MXT_CM-SUPn_tile.tif`) compositions of all input imagery with one file per band, `n` referring to the band number, the operations being restricted to the ROI extent considered during mosaicing (either DROI or CROI). The DN values of the pixels correspond to Top Of the Atmosphere (TOA) values rescaled between 0 and 250. The value 255 is reserved for no data value. These GeoTIFF files are of interest for visual quality control of the mosaic as advocated in [35]. For instance, a cloud appearing in the point-wise minimum composition indicates that this region is always covered by clouds so that they will also be visible in any mosaic produced with the same set of input imagery. Similarly, a region appearing as a cloud shadow in the point-wise minimum composition will always appear as such in a mosaic based on the same set of input imagery.

3. point-wise addition of cloud masks expanded in the direction opposite to the sun azimuth angle and restricted to the ROI extent considered during mosaicing (either DROI or CROI). Resulting aggregated cloud/shadow masks are stored int he file `date_MXT_CLSH_tile.tif`. The cloud masking algorithm is detailed in [35]. Beware that the North direction is pointing

[1] According to [40], it would be more precise to use the weighted median of the dates of the image pieces used for creating each tile.

upwards only for pixels centred on the meridian corresponding to the LAEA projection centre (i.e., 10^o East). The actual direction of the North of was calculated for the centre pixel of each scene and used for all pixels of the considered scene.

4. degree of overlap images (`date_MXT_GLVL_tile.tif`). This GeoTIFF file indicates the number of input images that is covering any given pixel. Hence, pixels set to 0 in this image are not covered by the input imagery. The maximum number of overlap levels is equal to 255 (maximum value for an unsigned byte integer).

6.2.2 Mosaic layers

The second category of layers holds the main outputs of the mosaicing procedure. It contains 4 main items:

1. mosaic of the input images given the produced decision regions, one file per band (`date_MXT_IP-BnOC_tile.tif`), `n` referring to the band number. The DN values of these GeoTIFF files correspond to Top Of the Atmosphere (TOA) values rescaled between 0 and 250. The value 255 is reserved for no data value.

2. labelled decision regions (`date_MXT_BASE_tile.tif`). These GeoTIFF files indicate from which image the value of each pixel should originate from. Indeed, there is a one-to-one correspondence between the label values and the input images. The correspondence is given in the ASCII file `tile.tif_list.txt` described hereafter (item number 3). The pixel data type used for storing the decision regions is an unsigned two byte integer. The maximum label value is equal to 65534, the value 65535 being reserved for no data. The labelled decision regions are the core of the mosaic since they indicate where the cut between overlapping images should take place and which part of the images should be thrown away;

3. lists of GeoTIFF files (`tile.tif_list.txt`). These two column ASCII files contain the correspondence table between labels of the decision regions and the file name of the updated ROI of the input images they correspond to. A typical line of these files looks like:

   49 22-B/20060723-1124_IL3_IM-UROI_2966422035-AC.tif

   That is, the image of the reference coverage uniquely defined by the date-time, sensor, and centre pixel fields of the indicated GeoTIFF file name should be used for all decision regions labelled with the value 49.

4. a series of look-up-tables (LUTs) mapping the label values to relevant information relative to their corresponding images. These LUTS are stored in ASCII files following the ENVI™ `dsr` (density slice range) format. The following LUTs are available:

   (a) sensor LUT mapping each label value to the sensor type of its corresponding image (`date_tile_sensor.dsr`). Sensors are represented with the following colours: red for SPOT-4, green for SPOT-5, and blue for IRS-LISS III.

(b) year LUT mapping each label value to the year of acquisition of the corresponding image (`date_tile_year.dsr`). The following colour codes are used: red for 2005, green for 2006, and blue for 2007;

(c) month LUT mapping each label value to the month of acquisition of its corresponding image (`date_tile_month.dsr`). Colour codes for months are derived from the colour wheel setting January to blue, February to cyan blue, Mars to cyan, April to cyan green, May to green, June to yellow green, August to orange, September to red, October to pink, November to magenta, and December to violet.

(d) week in year LUT mapping each label value to the week of acquisition of its corresponding image (`date_tile_week.dsr`). The colour wheel is also used for mapping a specific week to a colour, starting with blue for the first week and using violet for the last week of the year (colour wheel).

Because the label values are unique across all tiles, only one set of LUTs is saved for all tiles rather than storing subsets of these LUTs for each tile. The LUTs are therefore stored in a sub-directory named `ALL`, appearing at the same level as the tiles. This directory also holds a file containing the correspondence between label values and image names for all input imagery. It is named `UROI_list.txt`

6.2.3 Quality layers

The third category of layers contains files useful for a spatial representation of quantitative measurements of the geometric and radiometric consistency between adjacent decision regions of the mosaic. This is achieved by setting the boundary pixels of the decision regions to appropriate values. These values originate from the comprehensive relative and geometric consistency measurements detailed in Chap. 4. The quality layers contain 4 main items:

1–2. two images giving, for each pixel of the support of the morphological gradient [32] of the decision regions, the lowest (resp. greatest) label value in a 4-neighbourhood. These values are obtained by multiplying the support of the morphological gradient by the erosion (resp. dilation) of the labelled decision regions with a elementary diamond-shaped structuring element [33]. The images are stored in the files `date_MXT_BASE-ERO_tile-.tif` (resp. `date_MXT_BASE-DIL_tile.tif`).

3. a two dimensional LUT stored a TIFF image `date_ERODIL_LUT_HST_tile-.tif`. The value at coordinates x-y of this image indicates the number of pixels that have the value x in the image `date_MXT_BASE-ERO_tile.tif` and, at the same time, the value y in the image `date_MXT_BASE-DIL_tile-.tif`. It corresponds therefore to the two-dimensional frequency distribution using as input the 'erosion' and 'dilation' images described previously.

4. a series of additional two dimensional LUTs stored in the form of a TIFF image, each LUT summarising a consistency measurement. The following LUTs are available:

    1. `date_ERODIL_LUT_NORM_tile.tif` for the norm of the displacement vector between adjacent puzzle pieces;

2. `date_ERODIL_LUT_ADX_tile.tif` for the absolute value of the horizontal displacement between adjacent puzzle pieces;
3. `date_ERODIL_LUT_ADY_tile.tif` for the absolute value of the vertical displacement between adjacent puzzle pieces;
4. `date_ERODIL_LUT_COOR_tile.tif` for the radiometric correlation coefficient between the adjacent puzzle pieces.

For example, the value at coordinates x-y in the LUT `date_ERODIL_LUT_NORM_tile.tif` indicates that the norm of the mean displacement vector between the images indexed in the file `tile.tif_list.txt` by the values x and y. Similarly to the `dsr` LUTs described in Sec. 6.2.2, these 4 2-dimensional LUTs are stored in the directory `ALL`.

6.3 Coverage 1

Two mosaics products are available depending on whether data ROIs (DROIs) or their restriction to the country or countries the images were delivered for (CROIs) are used. The first mosaic maximises the use of the delivered imagery without taking into account the country/countries of origin of the input images. It provides a mosaic with as little gaps and clouds as possible but at the expense of geometric accuracy because the georeferencing specifications of an an image overlapping other countries than those it was delivered for are not necessarily met in these latter countries. The mosaic based on CROIs address this issue by considering only those parts of the image that fall within the country or the countries the image was delivered for (but at the expense of not using all available imagery).

In the Secs.6.3.1–6.3.3, results are presented according to the three categories described in Sec. 6.2 and using CROIs. Section 6.3.4 shows examples of differences obtained when considering DROIs instead of CROIs.

6.3.1 Base layers

Base layers described in Sec. 6.2.1 are not represented in this chapter Indeed, the degree of overlap images have already been illustrated for each coverage in Chap. 2 while examples of point-wise minimum and maximum compositions have been presented in Chap. 5. Examples of point-wise maximum compositions were also presented in Chap. 3. for a visual assessment of the quality of the proposed cloud detection algorithm.

6.3.2 Mosaic layers

Mosaic layers described in Sec. 6.2.2 are displayed in Figs. 6.1–6.6 in the following order: mosaic of multispectral imagery, labelled decision regions, and their mapping to sensor type as well as the year, month, and week of acquisition (thanks to the LUTs stored in the ASCII `dsr` files). Figure 6.7 and 6.8 shows a series of examples highlighting that seam lines largely follow salient image structures. The following colour codes are used for displaying increasing overlap levels: red (i.e., only one image available), green, blue, yellow, magenta, cyan, and white for overlap levels greater or equal to 0. It can be observed that

clouds occurring in the mosaic largely fall in regions where there is no overlap (see last 2 examples of Fig. 6.7). Note that, in some situations, there may be unfortunate image combinations leading to cut lines following ROI borders. Figure 6.9 illustrates 2 such cases. In the top case, there are too many overlapping images without clouds so that precedence was given to the few images not overlapping any other image (red regions in the image of overlaps). In the bottom case, an image contains so much clouds (bottom right section) so that it stops being used as soon as another image becomes available.

6.3.3 Quality layers

A direct graphical representation at European scale of the quality layers described in Sec. 6.2.3 is impossible given the scale of the figures presented in this document (2 pixel thick lines are not visible). Therefore, for visual purposes only, the measured quantities were expanded from the boundaries of the puzzle pieces using a dilation by an octagon of size 11 [33]. The resulting pictures obtained for the two main quality measures are displayed in Figs. 6.10 and 6.11: norm of the mean displacement vector and radiometric correlation coefficient. The majority of norms of the mean displacement vectors is less than half a pixel. However, in some areas, displacements between pairs of adjacent puzzle pieces well exceed this threshold value. For example, many neighbouring pieces in the Balkans reveal displacement vectors with a norm exceeding half or even a full pixel. It is also noticeable that most puzzle pieces in Iceland are separated by displacement vectors whose norms exceed one pixel. The correlation coefficient mapped in Fig. 6.11 have been calculated for the TOA values of the mid infrared band. One of the largest regions with correlation coefficient above 0.9 is the Iberian peninsula owing to a coverage with almost one sensor (IRS) and during more or less the same period of the year as revealed by Figs. 6.3 and 6.5 respectively.

6.3.4 From CROIs to DROIs

The same information layers are available for the mosaic of coverage 1 based on DROIs instead of CROIs. Figure 6.12 shows examples where the results obtained for DROIs depart from those obtained for CROIs. In some fewer cases, regions not covered by CROIs (i.e., imagery delivered for a country missing part of that country) is covered by imagery delivered by neighbouring countries and therefore are visible in the mosaic based on DROIs. This happens for instance for the Greek Megisti Island that is not included in the set of imagery delivered for Greece but appears in an image delivered for Turkey given its location near the coast of Turkey.

The norm of the displacement vector measured within the overlapping domains of the image pair meeting at the decision region boundaries and calculated on the basis of DROIs is mapped in Fig. 6.13 for all decision regions of the first coverage and bases on DROIs. The comparison of this figure with the one obtained for CROIs shows that larger displacements are sometimes obtained near the boundaries of the participating boundaries (see for example red line between Ireland and Northern Ireland occurring in Fig. 6.13 but not in Fig. 6.10. This was expected given the relative geometric consistency measurements for both CROIs and DROI summarised in Chap. 4.

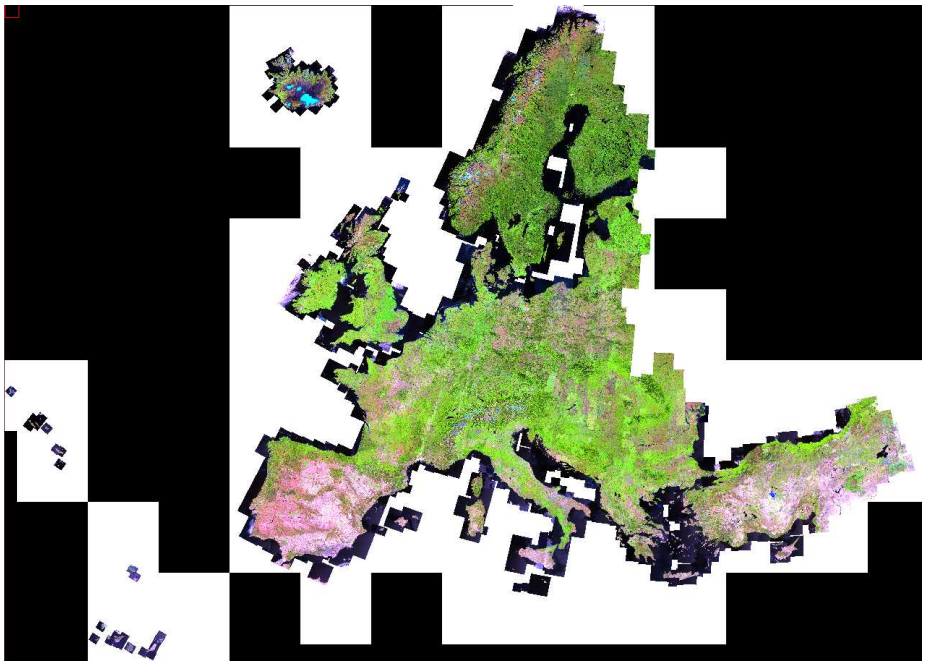


Figure 6.1: Top of atmosphere mosaic of first coverage based on CROIs and matching the labelled decision regions displayed in Fig. 6.2.

Figure 6.2: Labelled decision regions for the first coverage based on CROIs (1944 images used in mosaic out of 2004).

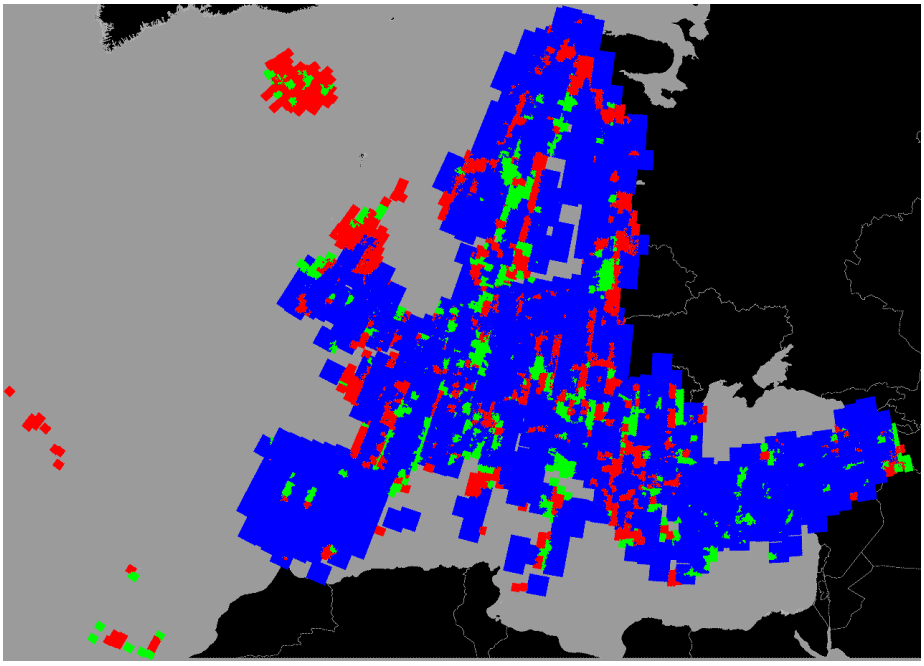


Figure 6.3: Decision regions of Fig. 6.2 mapped to sensor of image associated with each decision region of Fig. 6.2: red for SPOT-4, green for SPOT-5, blue for IRS-LISS.

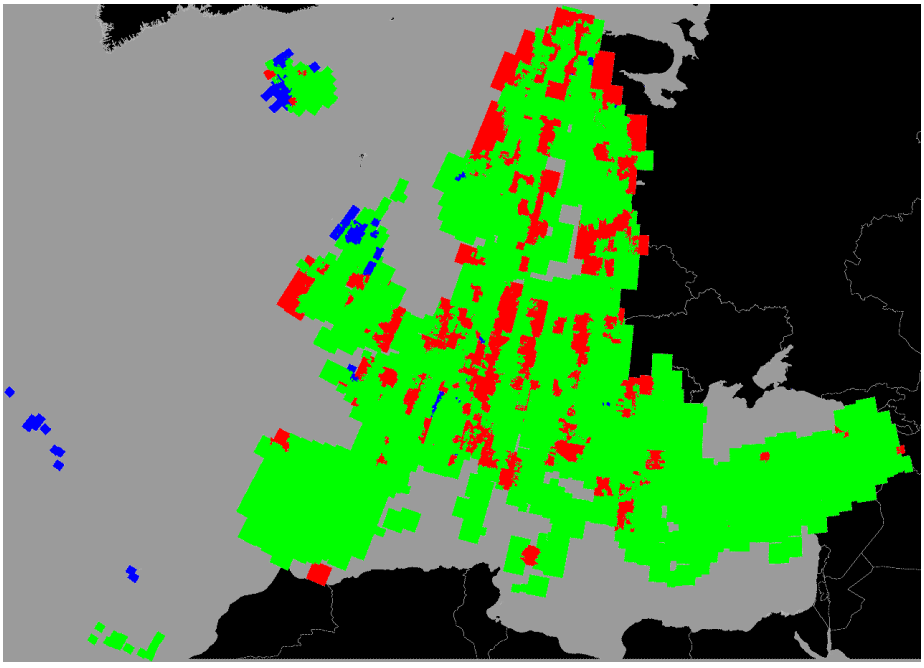


Figure 6.4: Decision regions of Fig. 6.2 mapped to year of acquisition of image associated with each decision region of Fig. 6.2: red for 2005, green for 2006, and blue for 2007.

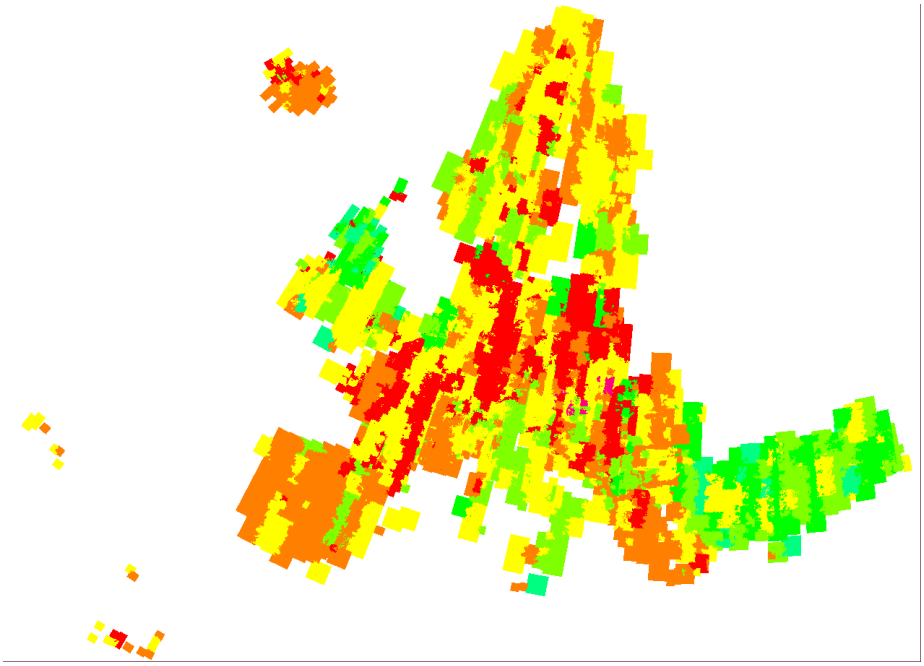


Figure 6.5: Decision regions of Fig. 6.2 mapped to month of acquisition of image associated with each decision region of Fig. 6.2. Month are coded using colour wheel, e.g., green for April, yellow for June, red for August, and magenta for October (see month LUT description on page 98 for the colour of each month).

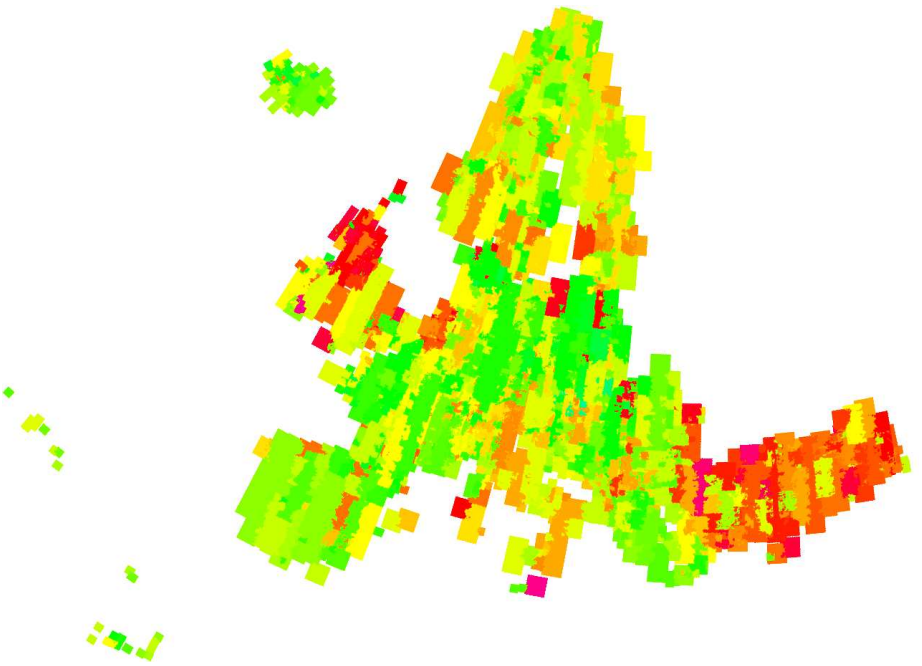


Figure 6.6: Decision regions of Fig. 6.2 mapped to week of acquisition of image associated with each decision region of Fig. 6.2.

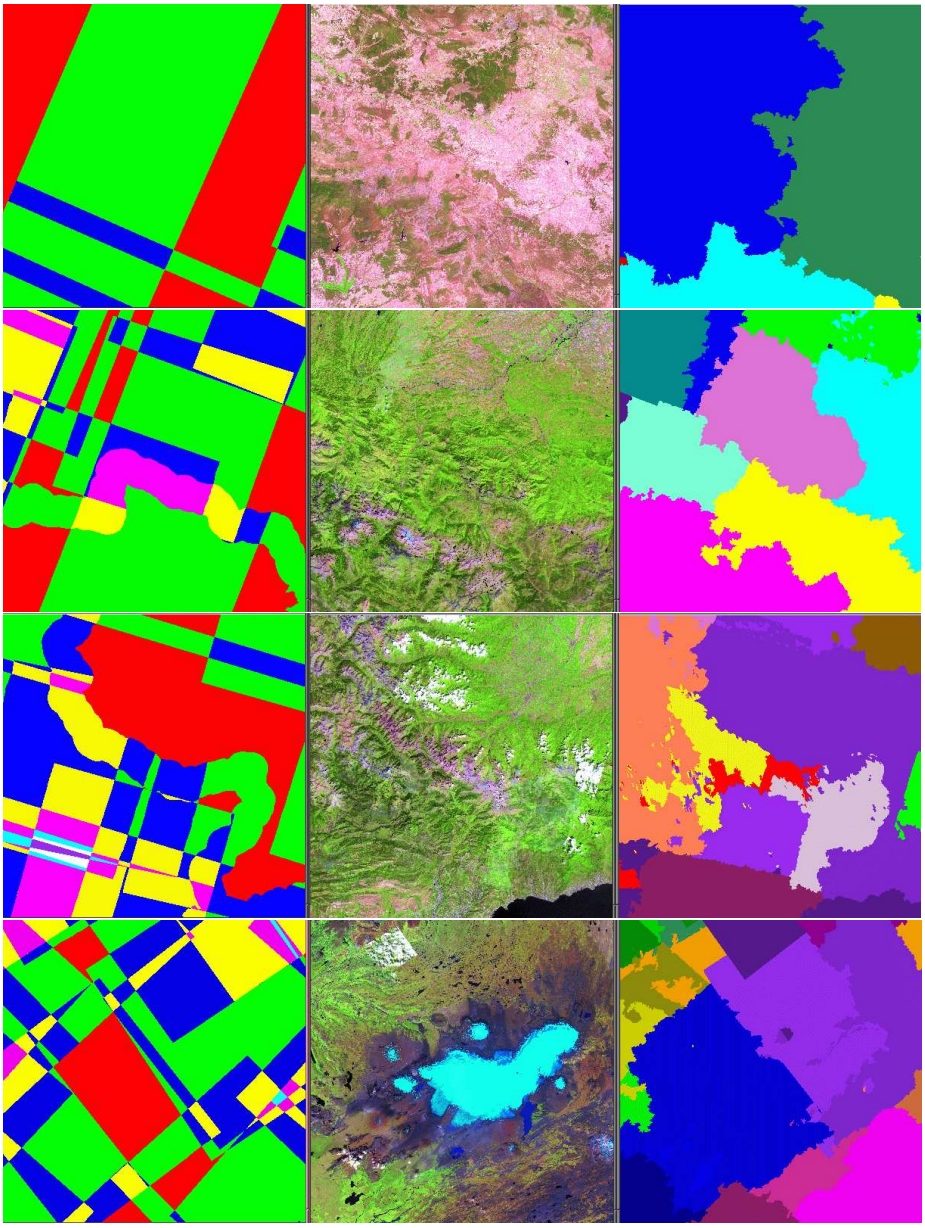


Figure 6.7: Gallery of overlap level images (left column), the resulting mosaic (middle column), and the underlying labelled decision regions (right column).

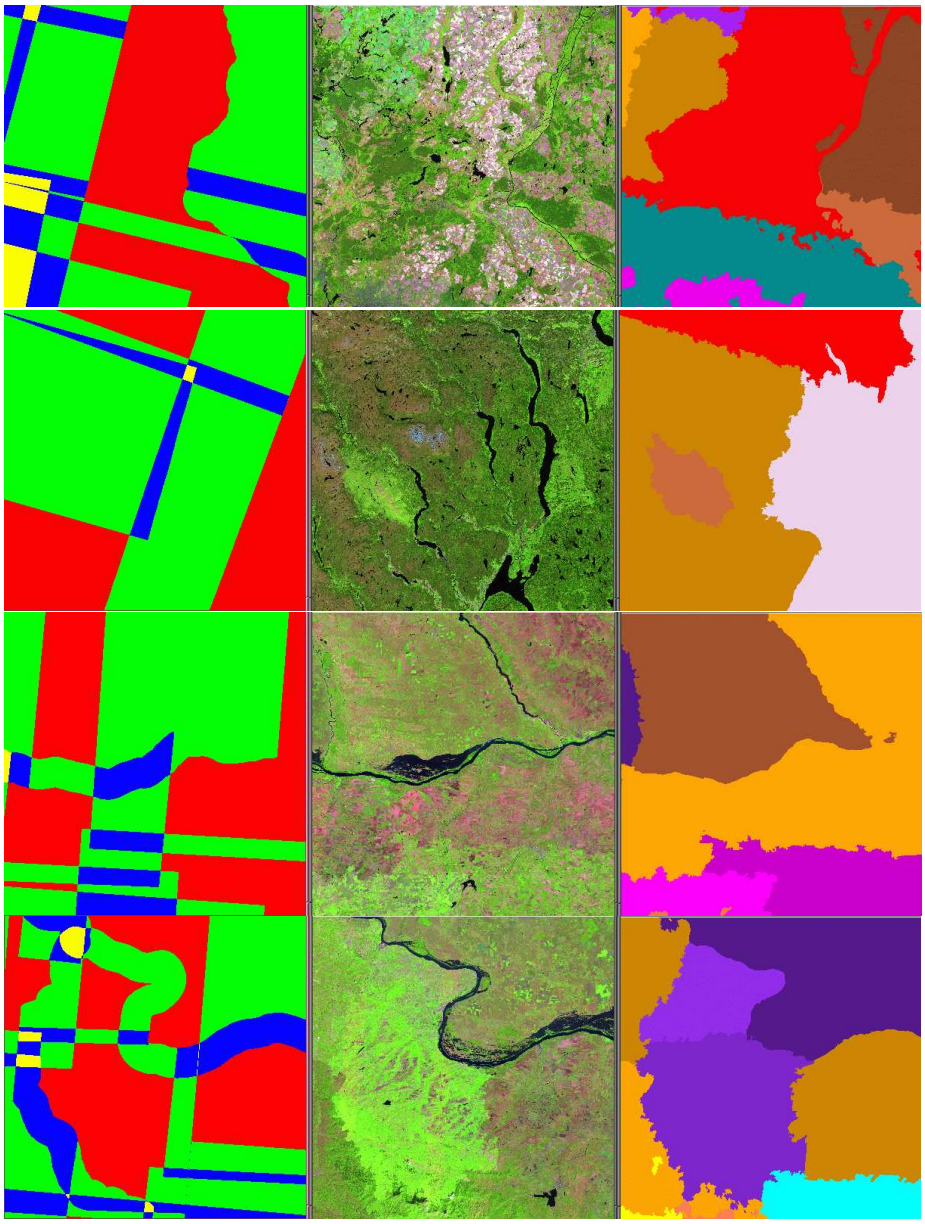


Figure 6.8: Gallery of overlap level images (left column), the resulting mosaic (middle column), and the underlying labelled decision regions (right column).

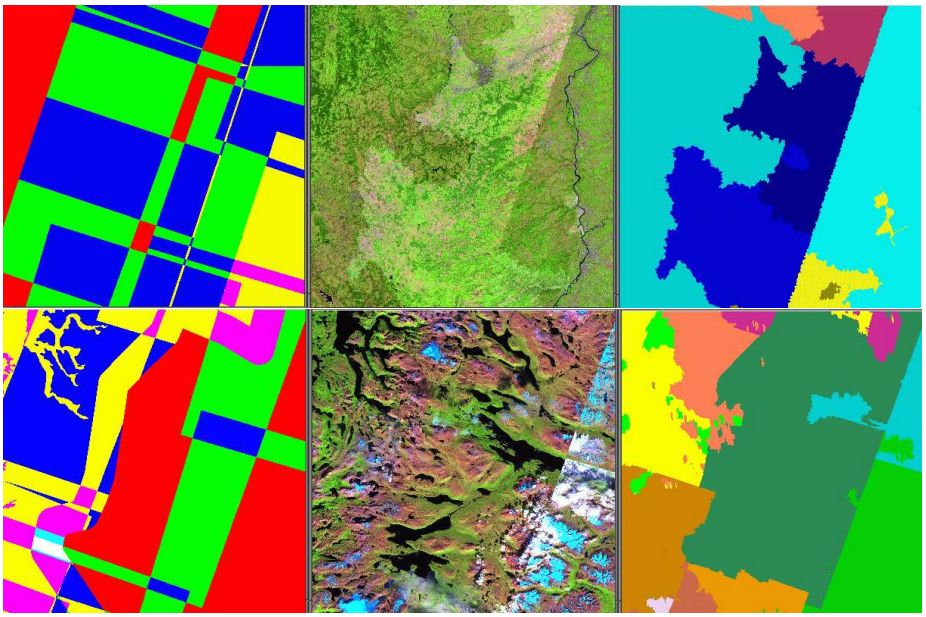


Figure 6.9: Gallery of overlap level images (left column), the resulting mosaic (middle column), and the underlying labelled decision regions (right column). In these examples, straight ROI borders are visible.

6.4 Coverage 2

The same comments as those made for coverage 1 applies. Results obtained for the mosaic layers based on DROIs are displayed in Figs. 6.15–6.19.

6.5 Conclusion

Morphological compositing [34] and its extension to the parallel order independent processing of large image data sets [36] enabled the automatic creation of a range of top-of-atmosphere reflectance mosaics minimising the visual detection of the cuts between overlapping images. Because morphological compositing permits the suppression of specific structures not appearing in at least one of the available images, clouds and their shadows were minimised thanks to their prior detection by the algorithm described in [35]. In summary, visually pleasing yet physically meaningful mosaics were created without applying cosmetic operations such as histogram balancing. That is, the values of the generated mosaics are comparable across the data set since they are given in TOA reflectance values. In addition, if necessary, it is possible to know from which image originates every single pixel of the mosaic. Beyond web mapping applications, all these properties make the mosaic with its accompanying information layers an added value product for subsequent analysis at the pan-European level.

Future research will aim at producing absolute rather than relative radiometric consistency measurements and, subsequently, map the TOA reflectance values to ground reflectance values by applying atmospheric correction procedures.

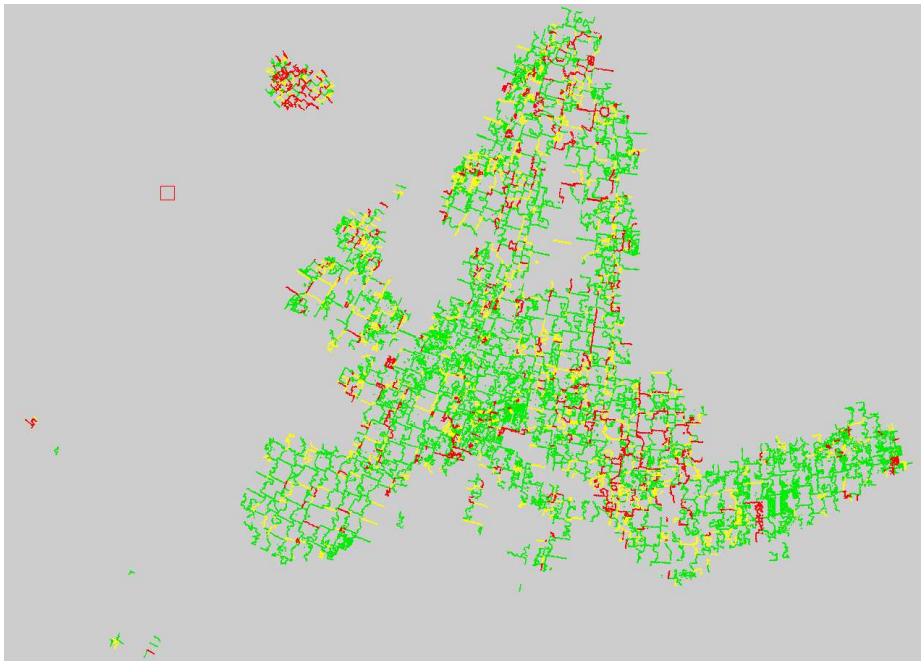


Figure 6.10: Norm d of the mean displacement vector measured within the overlapping domains of the image pair meeting at the decision region boundaries of Fig. 6.2. Green: $d \leq 0.5$ pixel. Yellow: $0.5 < d \leq 1.0$ pixel. Red: $d > 1$ pixel.

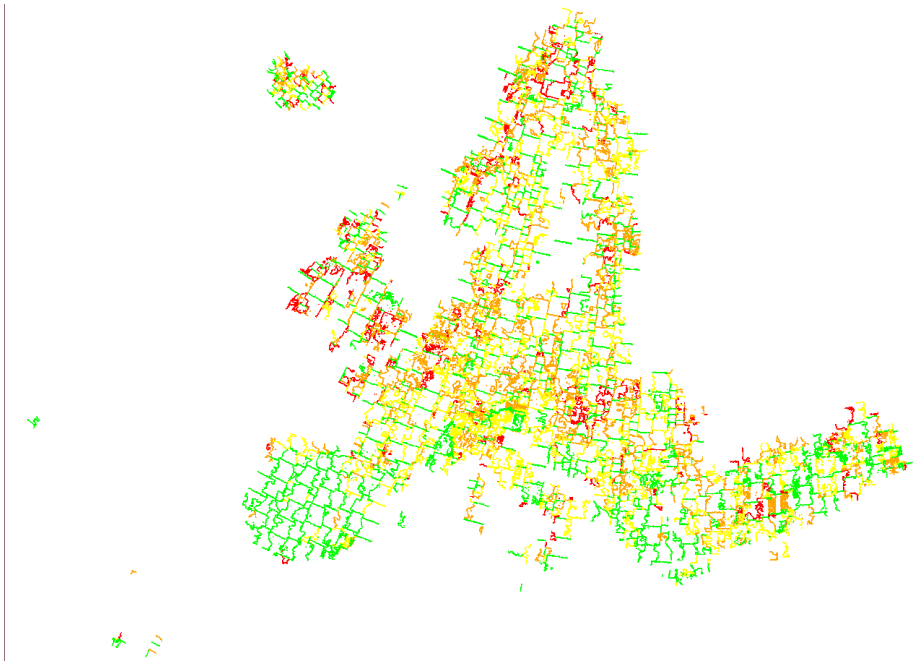


Figure 6.11: Radiometric correlation coefficient ρ measured within the overlapping domains of the image pair meeting at the decision region boundaries of Fig. 6.2. Green: $\rho > 0.9$. Yellow: $0.75 < \rho \leq 0.9$. Orange: $0.5 < \rho \leq 0.75$. Red: $\rho \leq 0.5$.

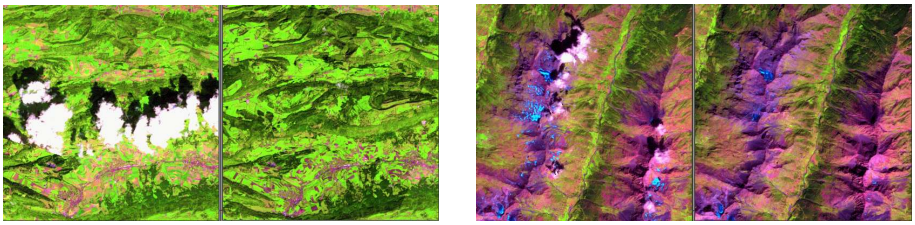


Figure 6.12: Coverage 1 mosaics: two examples of differences between CROI and DROI based mosaics. Right: near the border between Switzerland and France. Left: near the border between Italy and France.

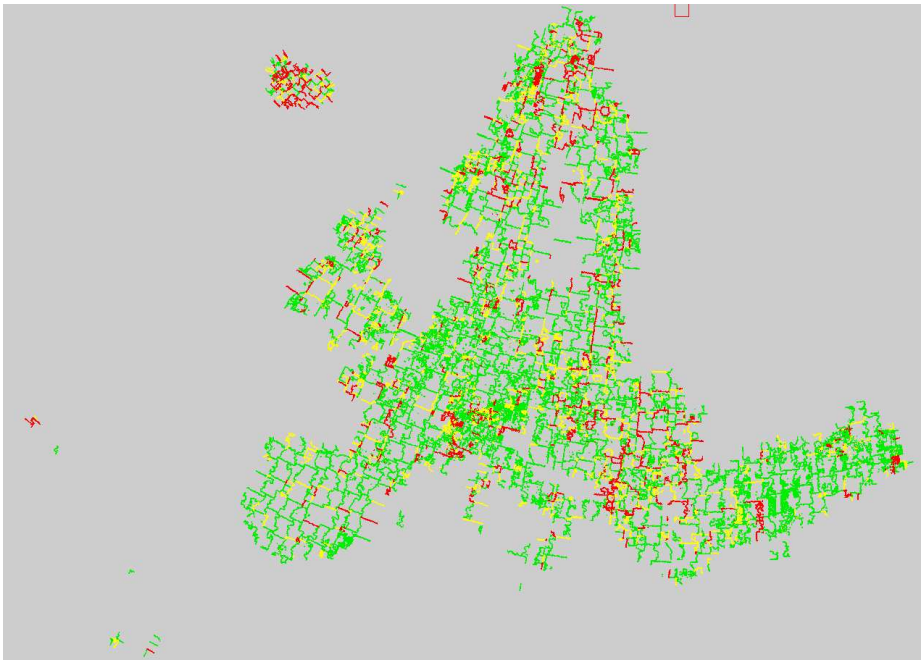


Figure 6.13: Norm d of the mean displacement vector measured within the overlapping domains of the image pair meeting at the DROI based decision region boundaries of the first coverage. Green: $d \leq 0.5$ pixel. Yellow: $0.5 < d \leq 1.0$ pixel. Red: $d > 1$ pixel. Compare with Fig. 6.10.

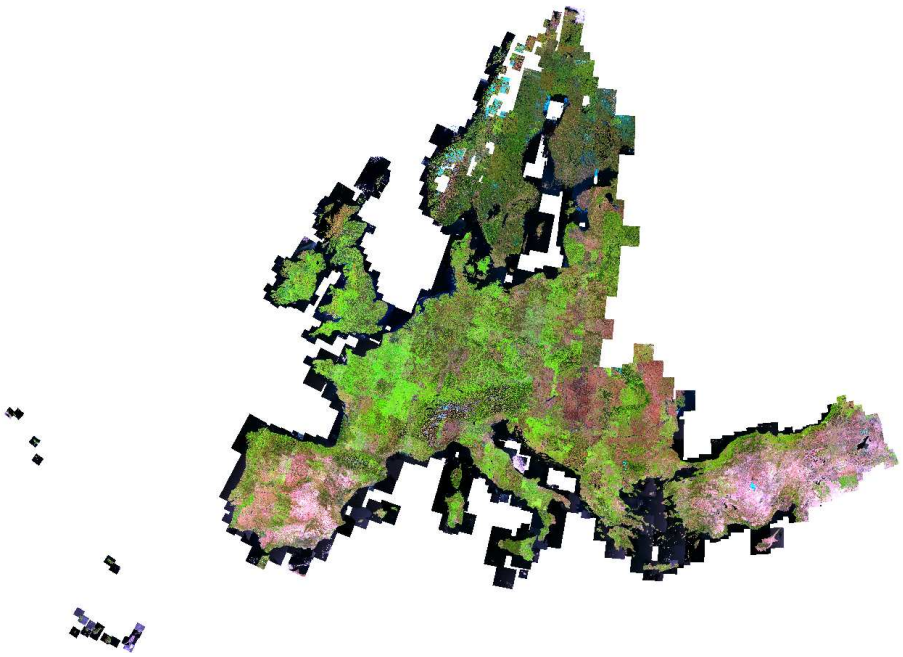


Figure 6.14: Top of atmosphere mosaic of second coverage based on DROIs and matching the labelled decision regions displayed in Fig. 6.15.

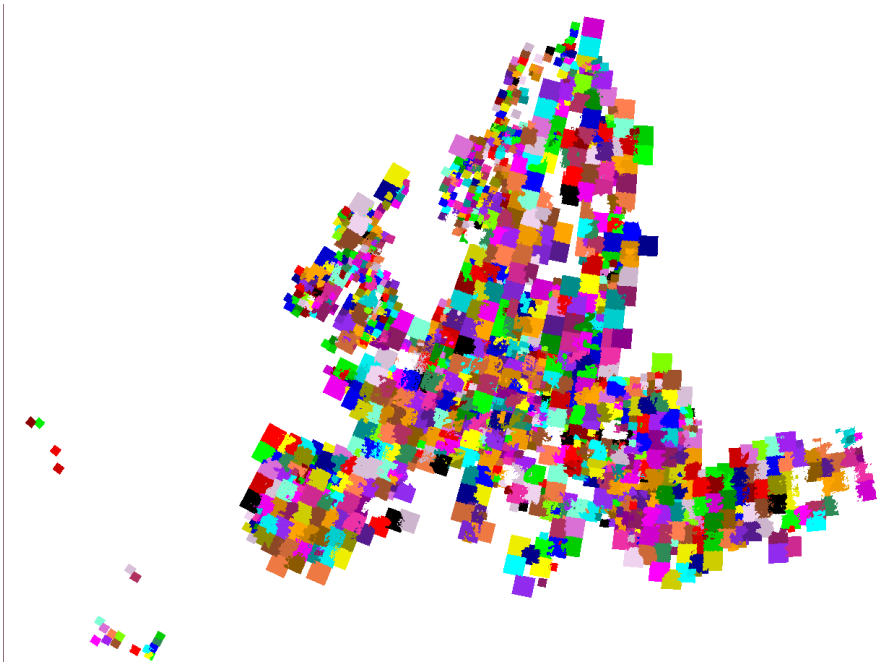


Figure 6.15: Labelled decision regions for the second coverage based on CROIs (1493 image used in mosaic out of 1561).

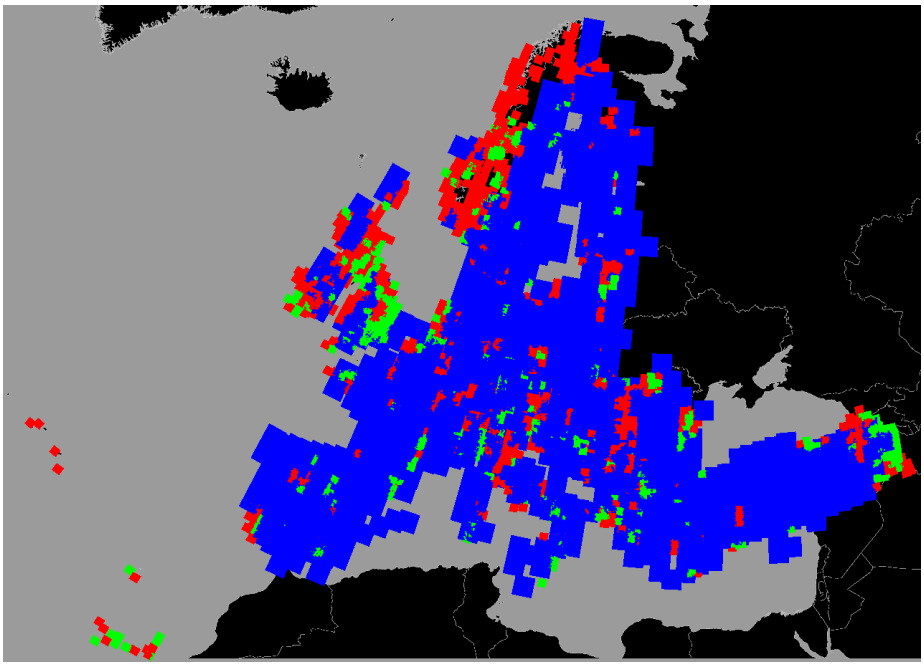


Figure 6.16: Decision regions of Fig. 6.15 mapped to sensor of image associated with each decision region of Fig. 6.15: red for SPOT-4, green for SPOT-5, blue for IRS-LISS.

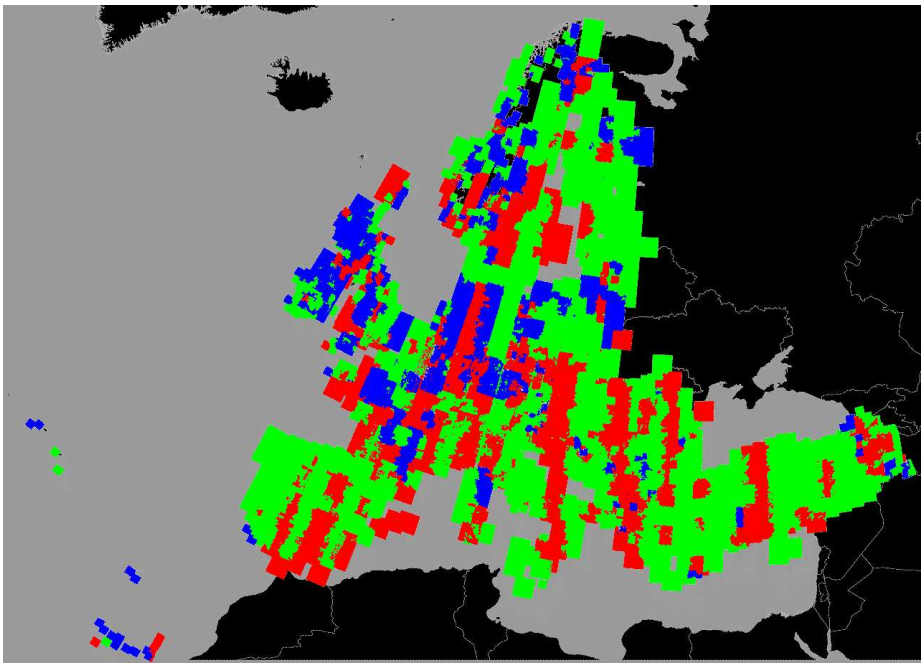


Figure 6.17: Decision regions of Fig. 6.15 mapped to year of acquisition of image associated with each decision region of Fig. 6.15: red for 2005, green for 2006, and blue for 2007.

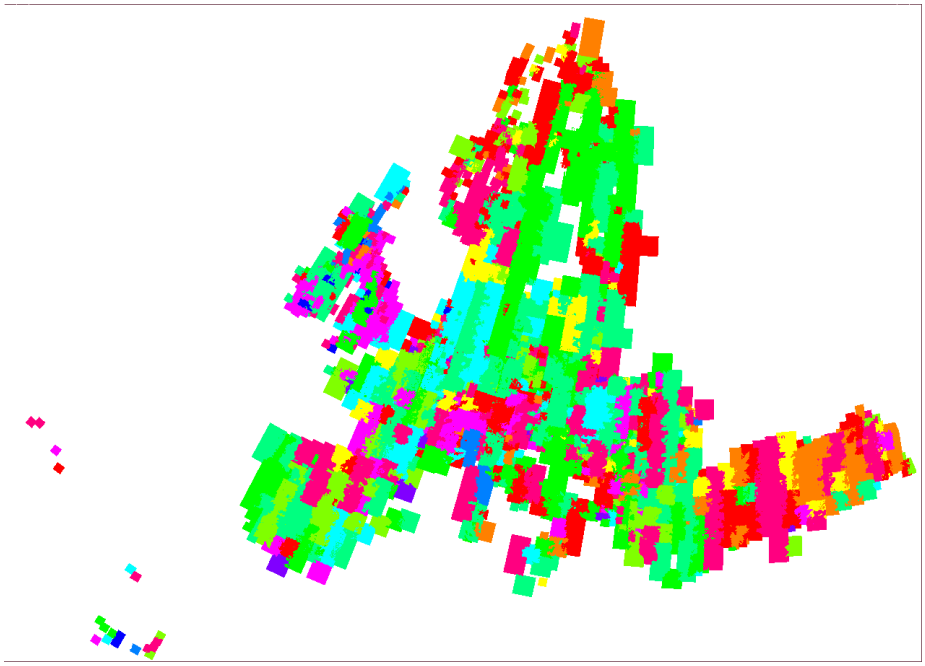


Figure 6.18: Decision regions of Fig. 6.15 mapped to month of acquisition of image associated with each decision region of Fig. 6.15. Month are coded using colour wheel, e.g., green for April, yellow for June, red for August, and magenta for October (see month LUT description on page 98 for the colour of each month).

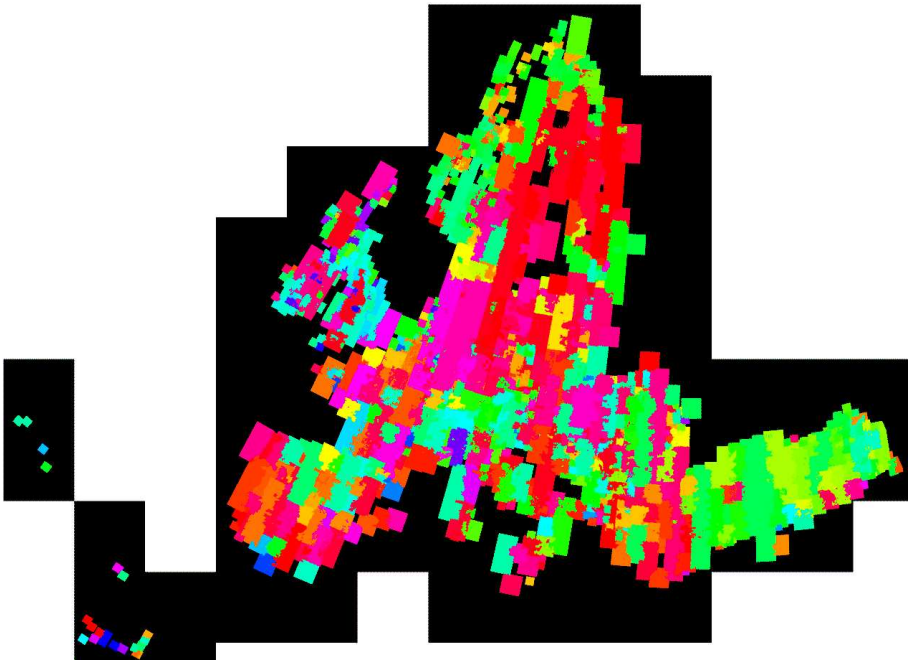


Figure 6.19: Decision regions of Fig. 6.15 mapped to week of acquisition of image associated with each decision region of Fig. 6.15.

Appendix A

Acronyms

ACCA ..... Automatic Cloud Cover Assessment
ASCII ..... American Standard Code for Information Interchange
COV1 ..... first coverage
COV2 ..... second coverage
CPU ..... Central Processing Unit
CROI ..... Country ROI (DROI restricted to given (buffered) country)
DLR Deutsches Zentrum für Lutf- und Raumfhart (German Aerospace Centre)
DN ..... Digital Number
DROI ..... Data ROI
ENVI™ ..... commercial software (ENvironement for Visualising Images)
ETRS89 ..... European Terrestrial Reference System 1989
GeoTIFF ..... Geocoded TIFF
GISCO ..... Geographic Information System of the European Commission
HRVIR ..... High Resolution Visible and Infrared
HRG ..... High Resolution Geometry
IL3 ..... IRS-LISS 3 sensor
IRS ..... Indian Remote Sensing Satellite (System)
ISO ..... International Organization for Standardization
LAEA ..... Lambert Azimuthal Equal Area
LISS ..... Linear Imaging Self Scanning Sensor
LUT ..... Look-Up-Table
LZW ..... Lempel Ziv Welsch
NUTS ..... National Territory Units
REF ..... REFerence coverage
RMSE ..... Root Mean Square Error
ROI ..... Region Of Interest
SPOT ..... Système Pour l'Observation de la Terre
SP4 ..... SPOT-4 HRVIR sensor
SP5 ..... SPOT-5 HRG sensor
TIFF ..... Tag Image File Format
TOA ..... Top Of Atmosphere
UROI ..... Updated ROI (decision region)
USGS ..... Unites States Geological Survey
`dsr` ..... file extension for ENVI™ ASCII density slice range files
`hdr` ..... file extension for ENVI™ ASCII header files

`tif` . . . . . file extension for TIFF or GeoTIFF images
`txt` . . . . . file extension for generic ASCII files

Appendix B

Country codes

List of two letter country codes used in the tables and the corresponding country names. Codes typeset with capital letters refer to ISO 3166-1-alpha-2 English country names and code elements [7] (those typeset in bold are referring to countries participating to the IMAGE-2006 project). This list does not represent the official position of the European instititutions with regard to the legal status or policy of the entities mentioned. It is a harmonisation of often divergent lists and practices. For more details, see `http://publications.europa.eu/code/en/en-5000500.htm`. Codes typeset with lower case letters are codes utilised in the data delivered to JRC and that deviates from the listed normalised codes.

AD ..... Andorra
AL ..... Albania
AM ..... Armenia
AT ..... Austria
AZ ..... Azerbaijan
BA ..... Bosnia and Herzegovina
BE ..... Belgium
BG ..... Bulgaria
BY ..... Belarus
CH ..... Switzerland
cs ..... Serbia
CS[1] ..... Former Union of Serbia (**RS**) and Montenegro (**ME**)
CY ..... Cyprus
CZ ..... Czech Republic
DE ..... Germany
DK ..... Denmark
EE ..... Estonia
ES ..... Spain
FI ..... Finland
FR ..... France
gb ..... Great Britain
GE ..... Georgia
GG ..... Guernsey

[1]This outdated code is used given that the available rasterised national territory units (GISCO NUTS level 0) contained only one polygon for Serbia and Montenegro.

GI ..... Gibraltar
GR ..... Greece
HR ..... Croatia
HU ..... Hungary
IE ..... Republic of Ireland
IM ..... Isle of Man
IQ ..... Iraq
IR ..... Iran
IS ..... Iceland
IT ..... Italy
JE ..... Jersey
LI ..... Lichtenstein
LT ..... Lituania
LU ..... Luxembourg
LV ..... Latvia
MA ..... Morocco
MC ..... Monaco
mc ..... the former Yugoslav Republic of Macedonia
MD ..... Moldova
ME ..... Montenegro
MK ..... the former Yugoslav Republic of Macedonia
MT ..... Malta
ni ..... Northern Ireland
NL ..... Netherlands
NO ..... Norway
PL ..... Poland
PT ..... Portugal
RO ..... Romania
RS ..... Serbia
RU ..... Russia
SE ..... Sweden
SI ..... Slovenia
SK ..... Slovakia
SM ..... San Marino
SY ..... Syria
TR ..... Turkey
UA ..... Ukraine
UK ..... United Kingdom
VA ..... Vatican City State (Holy See)

Appendix C

Comprehensive footprint coverage statistics

This appendix presents footprint statistics for all countries for which image is available. Tables are given for data and country based regions of interest and for coverages 1 (tables C.1 and C.2), coverage 2 (tables C.3 and C.4), as well as for the combined coverages (tables C.5 and C.6).

Table C.1: First coverage statistics based on data region of interest `DRXX`. The code CS refers to the former union of Serbia (RS) and Montenegro (ME). Country codes of participating countries are typeset in bold.

Code[a]	area[b] (km^2)	% with SP4	% with SP5	% with IL3	% with ALL	# missing pixels	Full coverage?
AD	465	0.00	0.00	100.00	100.00	0	yes
AL	28534	74.77	12.18	99.82	100.00	0	yes
AM	12130	0.12	3.26	39.64	41.19	27708600	no
AT	83930	51.47	27.66	92.59	100.00	0	yes
AZ	500	0.00	1.05	0.02	1.05	75777000	no
BA	51274	18.62	34.74	100.00	100.00	0	yes
BE	30664	28.15	63.79	100.00	≈100.00	100	no
BG	110796	61.25	27.10	100.00	100.00	0	yes
BY	30746	5.14	1.02	13.21	14.84	282372400	no
CH	41288	60.49	40.74	98.91	100.00	0	yes
CS	102572	45.29	31.28	93.25	100.00	0	yes
CY	9247	0.00	8.39	100.00	100.00	0	yes
CZ	78866	41.41	30.10	92.97	100.00	0	yes
DE	357672	34.76	39.99	96.88	100.00	0	yes
DK	43360	63.33	55.94	71.45	≈100.00	200	no
EE	45327	44.70	14.81	86.22	100.00	0	yes
ES	505984	6.96	12.01	95.28	100.00	0	yes
FI	337788	30.09	6.69	97.07	≈100.00	20700	no
FR	549164	39.99	23.65	92.24	≈100.00	3700	no
GE	21090	0.00	0.00	30.30	30.30	77610700	no
GG	8	10.08	0.00	0.00	10.08	118600	no
GI	5	0.00	0.00	100.00	100.00	0	yes

Code[a]	area[b] (km^2)	% with SP4	% with SP5	% with IL3	% with ALL	# missing pixels	Full coverage?
GR	132025	59.90	34.51	87.91	≈100.00	1300	no
HR	56631	18.22	32.96	99.62	≈100.00	500	no
HU	93016	49.89	50.51	86.91	100.00	0	yes
IE	70175	49.98	32.24	98.39	100.00	0	yes
IM	580	2.88	0.00	100.00	100.00	0	yes
IQ	6601	1.53	1.47	0.55	3.03	337589600	no
IR	10328	6.07	20.40	3.31	24.54	50806800	no
IS	102906	94.31	45.07	0.00	≈100.00	200	no
IT	301429	15.52	29.21	93.00	≈100.00	500	no
JE	123	23.56	0.00	100.00	100.00	0	yes
LI	160	100.00	0.00	100.00	100.00	0	yes
LT	64890	49.75	52.75	73.59	≈100.00	3600	no
LU	2595	37.15	77.78	100.00	100.00	0	yes
LV	64603	64.40	14.08	89.47	100.00	0	yes
MA	3734	0.02	0.00	1.03	1.03	576090200	no
MC	1	0.00	100.00	100.00	100.00	0	yes
MD	14423	0.52	2.54	40.18	42.60	31092900	no
MK	25153	55.58	0.88	100.00	100.00	0	yes
MT	314	0.00	0.00	100.00	100.00	0	yes
NL	37357	48.61	59.24	98.98	100.00	0	yes
NO	323456	41.05	6.59	98.47	≈100.00	100	no
PL	311894	42.47	8.68	98.14	100.00	0	yes
PT	92140	5.33	3.42	96.23	99.95	68500	no
RO	237938	32.49	32.87	93.42	100.00	0	yes
RU	112968	0.49	0.13	2.37	2.62	6731000100	no
SE	449446	29.95	40.04	87.02	≈100.00	2200	no
SI	20272	56.64	12.24	100.00	100.00	0	yes
SK	49024	64.86	34.42	96.31	100.00	0	yes
SM	61	0.00	0.00	100.00	100.00	0	yes
SY	35797	0.47	0.09	18.70	19.01	243984000	no
TR	780122	11.61	21.66	98.18	100.00	0	yes
UA	59237	0.76	0.26	9.85	9.86	866865400	no
UK	244706	55.11	17.83	83.86	99.99	26000	no
VA	0	0.00	0.00	100.00	100.00	0	yes

[a] Normalised country codes, all 38 participating countries are typeset in bold.
[b] Area of country calculated from Gisco version 9 NUTS vector file rasterised at 250m using ETRS-89 LAEA projection and the European grid and within window displayed in Fig. 2.1.

Table C.2: First coverage statistics based on country region of interest `CRXX`. The code CS refers to the former union of Serbia (RS) and Montenegro (ME). Country codes of participating countries are typeset in bold.

Code[a]	area[b] (km^2)	% with SP4	% with SP5	% with IL3	% with ALL	# missing pixels	Full coverage?
AD	465	0.00	0.00	100.00	100.00	0	yes
AL	28534	74.77	12.18	99.82	100.00	0	yes
AM	12130	0.12	3.26	39.64	41.19	27708600	no
AT	83930	51.47	27.66	92.59	100.00	0	yes
AZ	500	0.00	1.05	0.02	1.05	75777000	no
BA	51274	18.62	34.74	100.00	100.00	0	yes
BE	30664	28.15	63.79	100.00	≈100.00	100	no
BG	110796	61.25	27.10	100.00	100.00	0	yes
BY	30746	5.14	1.02	13.21	14.84	282372400	no
CH	41288	60.49	40.74	98.91	100.00	0	yes
CS	102572	45.29	31.28	93.25	100.00	0	yes
CY	9247	0.00	8.39	100.00	100.00	0	yes
CZ	78866	41.41	30.10	92.97	100.00	0	yes
DE	357672	34.76	39.99	96.88	100.00	0	yes
DK	43360	63.33	55.94	71.45	≈100.00	200	no
EE	45327	44.70	14.81	86.22	100.00	0	yes
ES	505984	6.96	12.01	95.28	100.00	0	yes
FI	337788	30.09	6.69	97.07	≈100.00	20700	no
FR	549164	39.99	23.65	92.24	≈100.00	3700	no
GE	21090	0.00	0.00	30.30	30.30	77610700	no
GG	8	10.08	0.00	0.00	10.08	118600	no
GI	5	0.00	0.00	100.00	100.00	0	yes
GR	132025	59.90	34.51	87.91	≈100.00	1300	no
HR	56631	18.22	32.96	99.62	≈100.00	500	no
HU	93016	49.89	50.51	86.91	100.00	0	yes
IE	70175	49.98	32.24	98.39	100.00	0	yes
IM	580	2.88	0.00	100.00	100.00	0	yes
IQ	6601	1.53	1.47	0.55	3.03	337589600	no
IR	10328	6.07	20.40	3.31	24.54	50806800	no
IS	102906	94.31	45.07	0.00	≈100.00	200	no
IT	301429	15.52	29.21	93.00	≈100.00	500	no
JE	123	23.56	0.00	100.00	100.00	0	yes
LI	160	100.00	0.00	100.00	100.00	0	yes
LT	64890	49.75	52.75	73.59	≈100.00	3600	no
LU	2595	37.15	77.78	100.00	100.00	0	yes
LV	64603	64.40	14.08	89.47	100.00	0	yes
MA	3734	0.02	0.00	1.03	1.03	576090200	no
MC	1	0.00	100.00	100.00	100.00	0	yes
MD	14423	0.52	2.54	40.18	42.60	31092900	no
MK	25153	55.58	0.88	100.00	100.00	0	yes
MT	314	0.00	0.00	100.00	100.00	0	yes
NL	37357	48.61	59.24	98.98	100.00	0	yes
NO	323456	41.05	6.59	98.47	≈100.00	100	no
PL	311894	42.47	8.68	98.14	100.00	0	yes

Code[a]	area[b] (km^2)	% with SP4	% with SP5	% with IL3	% with ALL	# missing pixels	Full coverage?
PT	92140	5.33	3.42	96.23	99.95	68500	no
RO	237938	32.49	32.87	93.42	100.00	0	yes
RU	112968	0.49	0.13	2.37	2.62	6731000100	no
SE	449446	29.95	40.04	87.02	≈100.00	2200	no
SI	20272	56.64	12.24	100.00	100.00	0	yes
SK	49024	64.86	34.42	96.31	100.00	0	yes
SM	61	0.00	0.00	100.00	100.00	0	yes
SY	35797	0.47	0.09	18.70	19.01	243984000	no
TR	780122	11.61	21.66	98.18	100.00	0	yes
UA	59237	0.76	0.26	9.85	9.86	866865400	no
UK	244706	55.11	17.83	83.86	99.99	26000	no
VA	0	0.00	0.00	100.00	100.00	0	yes

[a] Normalised country codes, all 38 participating countries are typeset in bold.
[b] Area of country calculated from Gisco version 9 NUTS vector file rasterised at 250m using ETRS-89 LAEA projection and the European grid and within window displayed in Fig. 2.1.

Table C.3: Second coverage statistics based on data region of interest `DRXX`. The code CS refers to the former union of Serbia (RS) and Montenegro (ME). Country codes of participating countries are typeset in bold.

Code[a]	area[b] (km^2)	% with SP4	% with SP5	% with IL3	% with ALL	# missing pixels	Full coverage?
AD	465	0.00	1.65	100.00	100.00	0	yes
AL	28534	26.83	23.91	96.93	100.00	0	yes
AM	5176	0.00	17.58	2.11	17.58	38834900	no
AT	83930	44.10	15.46	100.00	100.00	0	yes
AZ	822	0.00	1.72	0.00	1.72	75261300	no
BA	51274	12.95	34.82	98.82	100.00	0	yes
BE	30664	34.68	39.37	99.12	100.00	0	yes
BG	110796	17.42	20.19	97.24	100.00	0	yes
BY	50196	0.63	0.24	24.22	24.22	251252400	no
CH	41288	45.90	6.85	99.99	100.00	0	yes
CS	102572	40.95	12.13	98.50	100.00	0	yes
CY	9247	0.00	0.00	100.00	100.00	0	yes
CZ	78866	22.69	12.76	99.83	100.00	0	yes
DE	357672	15.25	7.86	99.65	100.00	0	yes
DK	43360	38.20	13.77	99.43	100.00	0	yes
EE	45327	0.00	8.00	100.00	100.00	0	yes
ES	505984	7.66	11.98	98.07	100.00	0	yes
FI	329895	23.49	3.66	90.49	97.66	12649900	no
FR	549167	11.77	6.25	97.81	100.00	0	yes
GE	6083	7.16	1.92	0.00	8.74	101622700	no
GI	5	0.00	0.00	100.00	100.00	0	yes
GR	132025	15.65	22.83	97.04	≈100.00	100	no
HR	56631	10.58	29.71	98.09	100.00	0	yes
HU	93016	26.95	3.70	97.06	100.00	0	yes
IE	70175	70.51	24.09	72.15	100.00	0	yes
IM	111	0.00	19.27	0.00	19.27	750400	no
IQ	9718	3.86	1.25	0.09	4.47	332601700	no
IR	10939	8.00	18.55	2.26	26.00	49828800	no
IT	301429	31.11	27.80	92.84	100.00	0	yes
JE	123	0.00	0.00	100.00	100.00	0	yes
LI	160	100.00	0.00	100.00	100.00	0	yes
LT	64892	44.22	16.49	100.00	100.00	0	yes
LU	2595	100.00	0.00	97.55	100.00	0	yes
LV	64603	13.99	4.27	99.96	100.00	0	yes
MA	7415	0.00	0.00	2.04	2.04	570200800	no
MC	1	0.00	0.00	100.00	100.00	0	yes
MD	20709	4.08	0.00	60.13	61.17	21036000	no
MK	25153	38.80	17.50	99.92	100.00	0	yes
MT	314	0.00	100.00	0.00	100.00	0	yes
NL	37357	69.17	39.92	72.21	100.00	0	yes
NO	274961	71.48	22.15	20.59	85.01	77592300	no
PL	311894	12.46	4.36	98.11	100.00	0	yes
PT	91223	36.15	11.80	89.30	98.96	1535000	no
RO	237938	37.98	12.73	87.44	100.00	0	yes

Code[a]	area[b] (km^2)	% with SP4	% with SP5	% with IL3	% with ALL	# missing pixels	Full coverage?
RU	118490	0.26	0.01	2.67	2.74	6722164800	no
SE	402998	15.87	13.36	78.71	89.67	74319400	no
SI	20272	44.60	33.99	96.61	100.00	0	yes
SK	49024	18.50	24.78	98.99	100.00	0	yes
SM	61	30.59	100.00	0.00	100.00	0	yes
SY	37994	1.76	0.47	18.56	20.18	240470000	no
TR	780122	9.20	9.56	93.68	100.00	0	yes
UA	44414	1.08	0.71	6.60	7.39	890581200	no
UK	244699	56.37	50.08	59.50	99.99	36100	no
VA	0	0.00	0.00	100.00	100.00	0	yes

[a] Normalised country codes, all 38 participating countries are typeset in bold.
[b] Area of country calculated from Gisco version 9 NUTS vector file rasterised at 250m using ETRS-89 LAEA projection and the European grid and within window displayed in Fig. 2.1.

Table C.4: Second coverage statistics based on country region of interest `CRXX`. The code CS refers to the former union of Serbia (RS) and Montenegro (ME). Country codes of participating countries are typeset in bold.

Code[a]	area[b] (km^2)	% with SP4	% with SP5	% with IL3	% with ALL	# missing pixels	Full coverage?
AD	465	0.00	1.49	100.00	100.00	0	yes
AL	28534	10.89	12.82	96.93	100.00	0	yes
AM	5176	0.00	17.58	2.11	17.58	38834900	no
AT	83930	37.44	15.42	100.00	100.00	0	yes
AZ	822	0.00	1.72	0.00	1.72	75261300	no
BA	51274	12.95	29.79	88.14	100.00	0	yes
BE	30664	29.73	26.99	80.86	100.00	0	yes
BG	110796	14.14	19.41	94.52	100.00	0	yes
BY	50196	0.63	0.24	24.22	24.22	251252400	no
CH	41288	41.81	1.64	96.91	100.00	0	yes
CS	102572	32.54	9.55	97.64	100.00	0	yes
CY	9247	0.00	0.00	100.00	100.00	0	yes
CZ	78866	22.69	12.14	98.25	100.00	0	yes
DE	357672	12.65	6.99	99.63	100.00	0	yes
DK	43360	38.20	13.77	94.07	100.00	0	yes
EE	45327	0.00	8.00	100.00	100.00	0	yes
ES	505984	6.58	12.07	98.07	100.00	0	yes
FI	329717	22.65	2.64	90.43	97.61	12933500	no
FR	549165	10.52	5.88	97.79	≈100.00	1900	no
GE	6083	7.16	1.92	0.00	8.74	101622700	no
GI	5	0.00	0.00	100.00	100.00	0	yes
GR	132025	13.29	22.10	97.04	≈100.00	100	no
HR	56631	8.14	16.19	95.97	100.00	0	yes
HU	93016	22.58	2.99	94.85	100.00	0	yes
IE	70175	63.34	19.25	72.15	100.00	0	yes
IM	111	0.00	19.27	0.00	19.27	750400	no
IQ	9718	3.86	1.25	0.09	4.47	332601700	no
IR	10939	8.00	18.55	2.26	26.00	49828800	no
IT	301429	30.73	27.74	92.59	100.00	0	yes
JE	123	0.00	0.00	100.00	100.00	0	yes
LI	160	100.00	0.00	100.00	100.00	0	yes
LT	64892	41.83	12.16	97.94	100.00	0	yes
LU	2595	100.00	0.00	35.81	100.00	0	yes
LV	64603	9.30	3.69	99.89	100.00	0	yes
MA	7415	0.00	0.00	2.04	2.04	570200800	no
MC	1	0.00	0.00	100.00	100.00	0	yes
MD	20709	4.08	0.00	60.13	61.17	21036000	no
MK	25153	19.09	8.34	99.76	100.00	0	yes
MT	314	0.00	100.00	0.00	100.00	0	yes
NL	37357	69.17	33.78	45.01	100.00	0	yes
NO	260664	69.63	22.15	17.27	80.59	100468000	no
PL	311894	11.37	3.69	97.15	100.00	0	yes
PT	91223	36.15	11.80	88.82	98.96	1535000	no
RO	237938	36.85	12.43	85.11	100.00	0	yes

Code[a]	area[b] (km^2)	% with SP4	% with SP5	% with IL3	% with ALL	# missing pixels	Full coverage?
RU	118489	0.26	0.01	2.67	2.74	6722167200	no
SE	393877	13.93	10.07	78.47	87.64	88913800	no
SI	20272	41.51	32.86	75.30	100.00	0	yes
SK	49024	12.08	16.68	98.91	100.00	0	yes
SM	61	30.59	100.00	0.00	100.00	0	yes
SY	37994	1.76	0.47	18.56	20.18	240470000	no
TR	780122	8.97	9.56	93.50	100.00	0	yes
UA	44414	1.08	0.71	6.60	7.39	890581200	no
UK	244699	56.37	50.08	58.01	99.99	36100	no
VA	0	0.00	0.00	100.00	100.00	0	yes

[a] Normalised country codes, all 38 participating countries are typeset in bold.
[b] Area of country calculated from Gisco version 9 NUTS vector file rasterised at 250m using ETRS-89 LAEA projection and the European grid and within window displayed in Fig. 2.1.

Table C.5: Second coverage statistics based on data region of interest `DRXX`. The code CS refers to the former union of Serbia (RS) and Montenegro (ME). Country codes of participating countries are typeset in bold.

Code[a]	area[b] (km^2)	% with SP4	% with SP5	% with IL3	% with ALL	# missing pixels	Full coverage?
AD	465	0.00	1.65	100.00	100.00	0	yes
AL	28534	84.82	29.71	99.84	100.00	0	yes
AM	12265	0.12	17.58	39.64	41.65	27491500	no
AT	83930	80.91	39.02	100.00	100.00	0	yes
AZ	822	0.00	1.72	0.02	1.72	75261300	no
BA	51274	31.43	57.89	100.00	100.00	0	yes
BE	30664	49.77	69.08	100.00	100.00	0	yes
BG	110796	64.94	43.06	100.00	100.00	0	yes
BY	50388	5.16	1.02	24.22	24.32	250946400	no
CH	41288	69.58	40.76	100.00	100.00	0	yes
CS	102572	63.57	39.42	100.00	100.00	0	yes
CY	9247	0.00	8.39	100.00	100.00	0	yes
CZ	78866	54.67	37.51	99.88	100.00	0	yes
DE	357672	45.34	44.01	100.00	100.00	0	yes
DK	43360	80.59	60.67	99.43	100.00	0	yes
EE	45327	44.70	22.81	100.00	100.00	0	yes
ES	505984	13.75	21.59	98.07	100.00	0	yes
FI	337801	39.15	9.11	100.00	100.00	0	yes
FR	549167	44.47	26.59	98.84	100.00	0	yes
GE	21090	7.16	1.92	30.30	30.30	77610700	no
GG	8	10.08	0.00	0.00	10.08	118600	no
GI	5	0.00	0.00	100.00	100.00	0	yes
GR	132025	62.45	44.91	98.69	100.00	0	yes
HR	56631	26.93	58.39	100.00	100.00	0	yes
HU	93016	60.19	53.82	98.71	100.00	0	yes
IE	70175	80.47	49.00	100.00	100.00	0	yes
IM	580	2.88	19.27	100.00	100.00	0	yes
IQ	10401	4.00	2.30	0.58	4.78	331509100	no
IR	12721	13.70	24.75	3.31	30.23	46977800	no
IS	102906	94.31	45.07	0.00	≈100.00	200	no
IT	301429	40.60	46.86	99.91	100.00	0	yes
JE	123	23.56	0.00	100.00	100.00	0	yes
LI	160	100.00	0.00	100.00	100.00	0	yes
LT	64892	70.15	55.67	100.00	100.00	0	yes
LU	2595	100.00	77.78	100.00	100.00	0	yes
LV	64603	69.14	14.83	100.00	100.00	0	yes
MA	7470	0.02	0.00	2.05	2.05	570112700	no
MC	1	0.00	100.00	100.00	100.00	0	yes
MD	21043	4.60	2.54	62.15	62.15	20501000	no
MK	25153	71.89	17.76	100.00	100.00	0	yes
MT	314	0.00	100.00	100.00	100.00	0	yes
NL	37357	83.64	76.53	100.00	100.00	0	yes
NO	323457	80.50	25.92	98.58	100.00	0	yes
PL	311894	49.42	10.56	100.00	100.00	0	yes

Code[a]	area[b] (km^2)	% with SP4	% with SP5	% with IL3	% with ALL	# missing pixels	Full coverage?
PT	92140	39.91	14.99	96.62	99.95	68500	no
RO	237938	59.52	42.84	99.61	100.00	0	yes
RU	143823	0.63	0.13	3.33	3.33	6681632500	no
SE	449446	40.58	45.35	99.18	≈100.00	2200	no
SI	20272	88.75	41.65	100.00	100.00	0	yes
SK	49024	66.69	56.05	100.00	100.00	0	yes
SM	61	30.59	100.00	100.00	100.00	0	yes
SY	40134	2.23	0.56	20.48	21.32	237044900	no
TR	780122	18.06	28.53	98.68	100.00	0	yes
UA	62545	1.64	0.92	10.23	10.41	861572300	no
UK	244714	80.26	58.21	92.16	≈100.00	13200	no
VA	0	0.00	0.00	100.00	100.00	0	yes

[a] Normalised country codes, all 38 participating countries are typeset in bold.
[b] Area of country calculated from Gisco version 9 NUTS vector file rasterised at 250m using ETRS-89 LAEA projection and the European grid and within window displayed in Fig. 2.1.

Table C.6: Second coverage statistics based on country region of interest `CRXX`. The code CS refers to the former union of Serbia (RS) and Montenegro (ME). Country codes of participating countries are typeset in bold.

Code[a]	area[b] (km^2)	% with SP4	% with SP5	% with IL3	% with ALL	# missing pixels	Full coverage?
AD	465	0.00	1.49	100.00	100.00	0	yes
AL	28534	68.42	17.32	99.84	100.00	0	yes
AM	12265	0.12	17.58	39.64	41.65	27491500	no
AT	83930	64.68	28.37	100.00	100.00	0	yes
AZ	822	0.00	1.72	0.02	1.72	75261300	no
BA	51274	24.59	43.98	100.00	100.00	0	yes
BE	30664	39.19	59.02	100.00	100.00	0	yes
BG	110796	60.06	42.92	97.86	100.00	0	yes
BY	50388	5.15	1.02	24.22	24.32	250946400	no
CH	41288	60.32	36.86	100.00	100.00	0	yes
CS	102572	58.42	32.73	100.00	100.00	0	yes
CY	9247	0.00	8.39	100.00	100.00	0	yes
CZ	78866	44.83	36.70	99.73	100.00	0	yes
DE	357672	40.59	40.83	100.00	100.00	0	yes
DK	43360	80.47	60.67	97.53	100.00	0	yes
EE	45327	44.67	22.81	100.00	100.00	0	yes
ES	505984	12.37	20.48	98.07	100.00	0	yes
FI	337801	37.61	8.06	100.00	100.00	0	yes
FR	549167	43.85	26.23	98.82	100.00	0	yes
GE	21090	7.16	1.92	30.30	30.30	77610700	no
GG	8	10.08	0.00	0.00	10.08	118600	no
GI	5	0.00	0.00	100.00	100.00	0	yes
GR	132025	61.13	41.59	98.04	100.00	0	yes
HR	56631	16.35	38.86	100.00	100.00	0	yes
HU	93016	54.28	49.91	97.98	100.00	0	yes
IE	70175	74.12	47.94	100.00	100.00	0	yes
IM	580	2.88	19.27	100.00	100.00	0	yes
IQ	10401	4.00	2.30	0.58	4.78	331509100	no
IR	12721	13.70	24.75	3.31	30.23	46977800	no
IS	102906	94.31	45.07	0.00	≈100.00	200	no
IT	301429	39.46	45.79	99.89	100.00	0	yes
JE	123	23.56	0.00	100.00	100.00	0	yes
LI	160	100.00	0.00	100.00	100.00	0	yes
LT	64892	61.17	55.37	98.01	100.00	0	yes
LU	2595	100.00	77.74	44.40	100.00	0	yes
LV	64603	65.66	13.14	100.00	100.00	0	yes
MA	7470	0.02	0.00	2.05	2.05	570112700	no
MC	1	0.00	100.00	100.00	100.00	0	yes
MD	21043	4.60	2.54	62.15	62.15	20501000	no
MK	25153	32.99	8.34	100.00	100.00	0	yes
MT	314	0.00	100.00	100.00	100.00	0	yes
NL	37357	82.22	76.52	98.81	100.00	0	yes
NO	323457	77.84	24.41	98.58	100.00	0	yes
PL	311894	46.56	9.80	99.21	100.00	0	yes

Code[a]	area[b] (km^2)	% with SP4	% with SP5	% with IL3	% with ALL	# missing pixels	Full coverage?
PT	92140	38.55	14.99	96.14	99.95	68500	no
RO	237938	56.74	40.32	99.03	100.00	0	yes
RU	143823	0.63	0.13	3.33	3.33	6681632500	no
SE	449446	39.66	43.79	96.71	≈100.00	2200	no
SI	20272	82.40	41.18	97.62	100.00	0	yes
SK	49024	52.50	47.13	100.00	100.00	0	yes
SM	61	30.59	100.00	100.00	100.00	0	yes
SY	40134	2.23	0.56	20.48	21.32	237044900	no
TR	780122	17.40	28.25	98.68	100.00	0	yes
UA	62545	1.64	0.92	10.23	10.41	861572300	no
UK	244714	80.26	58.21	92.15	≈100.00	13200	no
VA	0	0.00	0.00	100.00	100.00	0	yes

[a] Normalised country codes, all 38 participating countries are typeset in bold.
[b] Area of country calculated from Gisco version 9 NUTS vector file rasterised at 250m using ETRS-89 LAEA projection and the European grid and within window displayed in Fig. 2.1.

Appendix D

Comprehensive cloud statistics

This appendix lists the percentage of the covered territory that is always covered by clouds. All countries at least partly covered by the IMAGE-2006 coverages are considered in this table. The percentage of each country that is indeed covered by the IMAGE-2006 imagery is given in the appendix C. Using the latter table, it can be seen, for example, that the Isle of Man (IM) is fully covered in terms of DROIs in the first coverage but almost not covered in the second coverage. However, from the table in the present appendix, it can be seen that the Isle of Man is cluttered by 10.43% of clouds in the first coverage. By contrast, it can be seen that the Vatican City State (VA) is fully covered in all coverages (even when considering CROIs) and is not contaminated by any cloud.

Table D.1: Clouds remaining for all countries at least partly covered by the IMAGE-2006 data when considering available imagery for each coverage on the basis of DROIs and CROIs (the latter being restricted to buffered participating countries). In this table, percentages are relative to the portion of territory covered by imagery rather than the whole territory of the listed country. The ISO code of participating countries are typeset in bold (except for Serbia and Montenegro that are still appearing as the union of these two countries with the code CS).

	DROI			CROI		
	COV1	COV2	REF	COV1	COV2	REF
AD	0.00	2.08	0.00	0.00	2.16	0.00
AL	0.03	0.13	0.00	0.21	1.02	0.07
AM	5.96	1.46	4.97	5.96	1.46	4.97
AT	1.23	0.29	0.05	2.27	2.23	0.75
AZ	11.14	0.73	0.53	11.14	0.73	0.53
BA	0.00	0.08	0.00	0.43	1.06	0.01
BE	0.00	0.00	0.00	1.12	0.70	0.02
BG	0.03	0.48	0.00	0.28	0.75	0.00
BY	0.34	0.08	0.03	0.34	0.08	0.03
CH	0.80	0.22	0.07	5.31	4.81	0.70
CS	0.04	0.03	0.00	0.70	0.17	0.00

	DROI			CROI		
	COV1	COV2	REF	COV1	COV2	REF
CY	0.03	0.00	0.00	0.03	0.00	0.00
CZ	0.02	0.02	0.00	0.33	2.08	0.09
DE	0.06	0.02	0.00	0.29	0.28	0.00
DK	0.04	0.00	0.00	0.50	0.00	0.00
EE	0.04	0.02	0.00	0.04	0.02	0.00
ES	0.03	0.22	0.00	0.05	0.23	0.00
FI	0.03	0.30	0.00	0.20	0.31	0.02
FR	0.08	0.10	0.00	0.37	0.28	0.01
GE	8.80	0.15	8.59	8.80	0.15	8.59
GG	0.00	-	0.00	0.00	-	0.00
GI	0.00	0.00	0.00	0.00	0.00	0.00
GR	0.03	0.30	0.00	0.36	0.44	0.00
HR	0.04	0.03	0.00	0.36	0.12	0.01
HU	0.01	0.00	0.00	0.76	0.02	0.00
IE	0.26	0.23	0.00	1.78	0.73	0.00
IM	10.43	0.00	9.72	10.43	0.00	9.72
IQ	0.13	0.22	0.17	0.13	0.22	0.17
IR	2.22	0.90	0.81	2.22	0.90	0.81
IS	4.92	-	4.92	4.92	-	4.92
IT	0.32	0.13	0.01	0.90	0.27	0.04
JE	0.00	0.05	0.00	0.00	0.05	0.00
LI	0.00	0.00	0.00	0.46	0.15	0.15
LT	0.07	0.01	0.00	0.17	0.08	0.00
LU	0.00	0.00	0.00	1.20	2.38	0.73
LV	0.10	0.01	0.00	0.45	0.09	0.02
MA	0.03	4.37	3.48	0.03	4.37	3.48
MC	0.00	0.00	0.00	0.00	0.00	0.00
MD	0.03	0.28	0.02	0.03	0.28	0.02
MK	0.11	0.00	0.00	1.24	0.66	0.14
MT	0.89	0.03	0.00	0.89	0.03	0.00
NL	0.00	0.03	0.00	0.46	0.07	0.00
NO	0.74	1.61	0.18	1.23	1.72	0.36
PL	0.04	0.01	0.00	0.17	0.47	0.01
PT	0.03	0.19	0.03	0.03	0.19	0.03
RO	0.06	0.21	0.02	0.08	0.40	0.02
RU	0.05	1.14	0.34	0.05	1.14	0.34
SE	0.11	0.28	0.05	1.46	0.48	0.95
SI	0.01	0.11	0.00	2.30	3.43	0.32
SK	0.01	0.03	0.00	0.19	0.53	0.01
SM	0.00	0.60	0.00	0.00	0.60	0.00
SY	0.05	0.11	0.01	0.05	0.11	0.01
TR	0.32	0.18	0.02	0.33	0.25	0.02
UA	0.33	2.52	0.24	0.33	2.52	0.24
UK	0.56	0.33	0.00	0.59	0.34	0.00
VA	0.00	0.00	0.00	0.00	0.00	0.00

Appendix A

Between country geometric consistency

Displacement measurements between an image originating from a country and all those originating from neighbouring country are presented for each country hereafter (see tables A.1–A.38). Beware that all country codes used in these tables are referring to the codes used in the data delivered to JRC and therefore depart sometimes from normalised codes, see details in Chap 1 and appendix A.

Table A.1: Cross country geometric consistency assessment for **al**.

CC	#pair$_i$	$\overline{\lvert\overline{X}_i\rvert}$	$\overline{\lvert\overline{Y}_i\rvert}$	$\overline{\text{RMSE-}X_i}$	$\overline{\text{RMSE-}Y_i}$	$\lvert\overline{X}_i\rvert_{\max}$	$\lvert\overline{Y}_i\rvert_{\max}$
calculations based on DROIs							
ba	7	5.5	7.4	7.5	10.0	10.6	17.5
cs	64	5.6	4.0	8.4	6.8	27.3	12.7
gr	108	13.0	5.3	15.5	7.6	52.9	21.5
hr	4	2.7	5.9	4.2	7.4	5.0	11.5
it	8	4.3	7.1	9.9	12.3	7.8	14.1
mc	60	9.8	4.1	12.9	7.8	46.3	18.5
me	37	4.8	5.7	7.9	8.5	18.1	17.5
calculations based on CROIs							
cs	45	6.2	3.0	8.7	6.2	19.2	9.7
gr	66	7.6	5.0	11.1	8.7	41.7	28.0
it	8	3.8	7.0	9.9	13.9	8.1	13.9
mc	42	8.9	4.8	11.7	8.5	42.5	18.6
me	26	7.9	4.7	11.3	8.5	19.5	15.8

Table A.2: Cross country geometric consistency assessment for **at**.

CC	#pair$_i$	$\overline{\lvert\overline{X}_i\rvert}$	$\overline{\lvert\overline{Y}_i\rvert}$	$\overline{\text{RMSE-}X_i}$	$\overline{\text{RMSE-}Y_i}$	$\lvert\overline{X}_i\rvert_{\max}$	$\lvert\overline{Y}_i\rvert_{\max}$
calculations based on DROIs							
ba	2	10.5	4.2	12.3	6.9	14.3	5.0
ch	55	9.5	10.2	12.0	12.6	26.8	36.7
cz	159	3.9	5.9	6.9	8.7	16.7	23.8
de	188	4.2	4.5	7.2	7.4	21.9	25.7
hr	35	5.7	4.7	7.5	7.2	16.7	10.7
hu	84	6.5	6.4	10.0	10.7	24.2	26.1
it	123	7.4	5.5	10.2	9.0	31.7	21.8
li	11	3.3	1.6	5.9	4.3	11.0	4.6
pl	2	2.6	0.6	4.3	5.1	3.8	0.9
si	113	4.5	5.6	7.0	8.5	17.5	24.2
sk	48	6.2	4.0	11.1	9.9	23.6	15.7
calculations based on CROIs							
ch	32	8.7	7.5	11.1	10.3	26.4	25.3
cz	67	4.2	4.7	6.9	7.5	15.2	22.7
de	113	3.8	3.2	6.3	5.9	19.0	17.8
hr	3	21.6	16.7	27.2	21.4	28.2	21.9
hu	43	5.4	7.0	8.7	10.5	16.6	27.0
it	71	6.2	5.6	8.9	8.6	25.3	23.4
li	8	3.3	4.1	4.1	5.2	5.8	10.0
si	47	3.9	4.9	5.7	7.0	20.7	24.2
sk	16	3.2	8.2	6.2	11.4	13.6	25.0

Table A.3: Cross country geometric consistency assessment for **ba**.

CC	#pair$_i$	$\overline{\lvert\overline{X}_i\rvert}$	$\overline{\lvert\overline{Y}_i\rvert}$	$\overline{\text{RMSE-}X_i}$	$\overline{\text{RMSE-}Y_i}$	$\lvert\overline{X}_i\rvert_{\max}$	$\lvert\overline{Y}_i\rvert_{\max}$
calculations based on DROIs							
al	7	5.5	7.4	7.5	10.0	10.6	17.5
at	2	10.5	4.2	12.3	6.9	14.3	5.0
cs	95	3.8	7.8	9.2	12.2	20.6	40.6
hr	207	4.5	5.4	7.8	8.6	28.4	40.6
hu	13	8.7	12.6	16.3	18.0	22.5	32.1
me	48	4.4	5.6	6.6	7.9	20.5	14.9
ro	1	0.0	35.9	8.9	37.1	0.0	35.9
si	10	17.5	7.7	20.0	11.1	47.0	21.1
calculations based on CROIs							
cs	50	3.6	6.4	9.6	11.3	16.0	21.8
hr	151	4.4	5.2	8.5	8.7	27.3	29.1
me	36	4.1	5.7	6.5	8.2	17.2	18.7

Table A.4: Cross country geometric consistency assessment for **be**.

CC	#pair$_i$	$\overline{\lvert X_i \rvert}$	$\overline{\lvert Y_i \rvert}$	$\overline{\text{RMSE-}X_i}$	$\overline{\text{RMSE-}Y_i}$	$\lvert X_i \rvert_{\max}$	$\lvert Y_i \rvert_{\max}$
calculations based on DROIs							
de	86	3.0	6.6	5.3	9.4	19.0	28.1
fr	134	4.0	5.2	7.5	8.8	21.3	26.1
lu	38	7.5	5.2	9.3	7.7	28.2	18.0
nl	151	3.0	3.5	6.9	7.7	13.6	25.1
calculations based on CROIs							
de	26	2.7	7.8	4.2	9.2	8.9	19.8
fr	75	4.7	6.5	9.1	10.6	21.0	32.2
lu	19	10.2	3.2	11.4	5.6	29.8	9.7
nl	92	3.0	3.3	8.0	8.6	15.2	15.4

Table A.5: Cross country geometric consistency assessment for **bg**.

CC	#pair$_i$	$\overline{\lvert X_i \rvert}$	$\overline{\lvert Y_i \rvert}$	$\overline{\text{RMSE-}X_i}$	$\overline{\text{RMSE-}Y_i}$	$\lvert X_i \rvert_{\max}$	$\lvert Y_i \rvert_{\max}$
calculations based on DROIs							
cs	121	10.7	7.2	13.1	9.8	35.8	26.8
gr	176	6.3	7.5	9.2	9.9	27.3	26.5
mc	70	6.4	5.9	8.8	8.2	28.4	23.7
ro	194	8.7	8.2	16.5	15.1	54.8	39.9
tr	64	8.1	5.1	11.1	8.3	26.3	14.5
calculations based on CROIs							
cs	69	10.6	7.7	12.3	9.7	29.7	25.8
gr	81	6.8	5.3	8.6	7.5	27.2	26.2
mc	41	4.1	3.5	5.7	5.9	12.6	16.7
ro	97	8.5	8.0	20.8	17.7	50.7	37.6
tr	36	6.3	5.7	8.0	7.3	20.4	17.9

Table A.6: Cross country geometric consistency assessment for **ch**.

CC	#pair$_i$	$\overline{\lvert X_i \rvert}$	$\overline{\lvert Y_i \rvert}$	$\overline{\text{RMSE-}X_i}$	$\overline{\text{RMSE-}Y_i}$	$\lvert X_i \rvert_{\max}$	$\lvert Y_i \rvert_{\max}$
calculations based on DROIs							
at	55	9.5	10.2	12.0	12.6	26.8	36.7
de	125	7.6	7.8	10.0	10.9	25.8	36.7
fr	82	7.8	9.3	10.3	12.4	30.1	33.4
it	173	8.8	13.0	13.0	16.3	31.7	41.4
li	29	9.9	8.5	11.7	11.2	23.5	34.0
calculations based on CROIs							
at	32	8.7	7.5	11.1	10.3	26.4	25.3
de	47	4.8	5.1	8.8	9.3	13.8	18.3
fr	69	7.8	9.5	10.0	12.8	30.4	36.2
it	130	6.5	10.2	11.4	13.7	19.3	33.8
li	10	8.2	9.3	10.1	11.1	17.5	25.5

Table A.7: Cross country geometric consistency assessment for **cs**.

CC	#pair$_i$	$\overline{\lvert\overline{X}_i\rvert}$	$\overline{\lvert\overline{Y}_i\rvert}$	$\overline{\text{RMSE-}X_i}$	$\overline{\text{RMSE-}Y_i}$	$\lvert\overline{X}_i\rvert_{\max}$	$\lvert\overline{Y}_i\rvert_{\max}$
calculations based on DROIs							
al	64	5.6	4.0	8.4	6.8	27.3	12.7
ba	95	3.8	7.8	9.2	12.2	20.6	40.6
bg	121	10.7	7.2	13.1	9.8	35.8	26.8
gr	5	6.0	14.6	7.9	16.6	9.2	19.9
hr	47	6.0	8.8	12.3	15.2	27.5	53.4
hu	69	7.3	6.2	12.8	11.7	27.7	18.2
mc	62	6.8	5.2	10.0	8.3	28.4	23.7
me	78	4.4	4.9	6.7	7.3	18.7	14.6
ro	142	9.0	7.1	13.1	11.4	26.8	50.7
calculations based on CROIs							
al	45	6.2	3.0	8.7	6.2	19.2	9.7
ba	50	3.6	6.4	9.6	11.3	16.0	21.8
bg	69	10.6	7.7	12.3	9.7	29.7	25.8
hr	36	5.2	11.2	11.5	16.0	19.9	53.2
hu	28	6.7	5.3	10.2	9.6	22.5	13.1
mc	38	6.8	4.9	9.1	7.5	26.1	22.4
me	76	4.4	4.8	6.6	7.1	20.9	14.6
ro	66	7.4	5.4	12.5	10.9	24.8	19.7

Table A.8: Cross country geometric consistency assessment for **cy**.

CC	#pair$_i$	$\overline{\lvert\overline{X}_i\rvert}$	$\overline{\lvert\overline{Y}_i\rvert}$	$\overline{\text{RMSE-}X_i}$	$\overline{\text{RMSE-}Y_i}$	$\lvert\overline{X}_i\rvert_{\max}$	$\lvert\overline{Y}_i\rvert_{\max}$
calculations based on DROIs							
tr	1	23.9	7.2	23.9	8.0	23.9	7.2
calculations based on CROIs							

Table A.9: Cross country geometric consistency assessment for **cz**.

CC	#pair$_i$	$\overline{\lvert\overline{X}_i\rvert}$	$\overline{\lvert\overline{Y}_i\rvert}$	$\overline{\text{RMSE-}X_i}$	$\overline{\text{RMSE-}Y_i}$	$\lvert\overline{X}_i\rvert_{\max}$	$\lvert\overline{Y}_i\rvert_{\max}$
calculations based on DROIs							
at	159	3.9	5.9	6.9	8.7	16.7	23.8
de	179	4.0	4.4	6.5	7.3	21.8	23.8
hu	14	6.0	7.1	12.9	12.4	15.1	16.3
pl	176	4.8	3.7	7.7	6.7	28.3	17.1
sk	53	4.4	2.2	8.6	6.8	20.7	8.3
calculations based on CROIs							
at	67	4.2	4.7	6.9	7.5	15.2	22.7
de	121	3.5	4.5	5.2	6.3	15.4	17.3
pl	85	4.2	3.5	6.5	6.1	27.0	20.4
sk	32	5.4	2.3	6.8	4.1	18.8	5.4

Table A.10: Cross country geometric consistency assessment for **de**.

CC	#pair$_i$	$\overline{\lvert X_i \rvert}$	$\overline{\lvert Y_i \rvert}$	$\overline{\text{RMSE-}X_i}$	$\overline{\text{RMSE-}Y_i}$	$\lvert X_i \rvert_{\max}$	$\lvert Y_i \rvert_{\max}$
calculations based on DROIs							
at	188	4.2	4.5	7.2	7.4	21.9	25.7
be	86	3.0	6.6	5.3	9.4	19.0	28.1
ch	125	7.6	7.8	10.0	10.9	25.8	36.7
cz	179	4.0	4.4	6.5	7.3	21.8	23.8
dk	138	4.7	4.9	12.3	12.6	23.5	22.9
fr	214	4.5	6.4	7.0	9.1	18.7	33.2
it	51	8.9	10.5	12.4	12.8	26.6	49.5
li	23	4.0	8.6	6.6	10.7	16.4	27.1
lu	42	4.0	5.8	7.0	8.8	18.1	21.4
nl	205	3.6	4.7	8.6	9.4	19.7	34.5
pl	116	4.1	4.6	8.1	8.2	22.0	22.0
se	2	5.1	17.6	8.8	21.3	9.3	17.6
si	1	9.5	12.8	12.5	16.5	9.5	12.8
calculations based on CROIs							
at	113	3.8	3.2	6.3	5.9	19.0	17.8
be	26	2.7	7.8	4.2	9.2	8.9	19.8
ch	47	4.8	5.1	8.8	9.3	13.8	18.3
cz	121	3.5	4.5	5.2	6.3	15.4	17.3
dk	54	4.7	6.5	16.2	15.7	22.7	17.8
fr	118	4.5	5.1	6.9	8.7	20.2	23.6
lu	28	4.2	5.6	6.4	8.3	18.2	17.8
nl	118	4.1	4.4	11.1	11.4	39.8	21.7
pl	60	4.3	4.2	7.7	7.3	26.1	17.7
se	1	1.1	12.3	7.7	30.2	1.1	12.3

Table A.11: Cross country geometric consistency assessment for **dk**.

CC	#pair$_i$	$\overline{\lvert X_i \rvert}$	$\overline{\lvert Y_i \rvert}$	$\overline{\text{RMSE-}X_i}$	$\overline{\text{RMSE-}Y_i}$	$\lvert X_i \rvert_{\max}$	$\lvert Y_i \rvert_{\max}$
calculations based on DROIs							
de	138	4.7	4.9	12.3	12.6	23.5	22.9
no	4	5.2	19.3	14.4	20.7	7.9	22.1
se	68	7.5	9.1	13.8	14.2	48.7	41.8
calculations based on CROIs							
de	54	4.7	6.5	16.2	15.7	22.7	17.8
se	22	6.2	11.1	14.3	17.1	27.2	26.3

Table A.12: Cross country geometric consistency assessment for **ee**.

CC	#pair$_i$	$\overline{\lvert X_i \rvert}$	$\overline{\lvert Y_i \rvert}$	$\overline{\text{RMSE-}X_i}$	$\overline{\text{RMSE-}Y_i}$	$\lvert \overline{X}_i \rvert_{\max}$	$\lvert \overline{Y}_i \rvert_{\max}$
calculations based on DROIs							
fi	31	12.8	9.4	14.8	11.9	34.5	24.1
lt	3	4.4	7.8	6.9	9.2	9.3	8.4
lv	68	3.8	3.6	6.4	6.1	19.3	12.2
calculations based on CROIs							
fi	14	8.7	5.6	10.4	6.9	25.8	20.3
lv	42	3.7	3.8	6.9	7.2	19.7	13.2

Table A.13: Cross country geometric consistency assessment for **es**.

CC	#pair$_i$	$\overline{\lvert X_i \rvert}$	$\overline{\lvert Y_i \rvert}$	$\overline{\text{RMSE-}X_i}$	$\overline{\text{RMSE-}Y_i}$	$\lvert \overline{X}_i \rvert_{\max}$	$\lvert \overline{Y}_i \rvert_{\max}$
calculations based on DROIs							
fr	208	3.7	6.2	6.7	9.2	23.3	25.3
pt	219	4.1	3.3	6.3	5.6	21.7	19.1
calculations based on CROIs							
fr	120	2.9	6.3	6.0	9.1	15.2	24.9
pt	133	4.0	2.9	5.1	4.5	17.8	16.7

Table A.14: Cross country geometric consistency assessment for **fi**.

CC	#pair$_i$	$\overline{\lvert X_i \rvert}$	$\overline{\lvert Y_i \rvert}$	$\overline{\text{RMSE-}X_i}$	$\overline{\text{RMSE-}Y_i}$	$\lvert \overline{X}_i \rvert_{\max}$	$\lvert \overline{Y}_i \rvert_{\max}$
calculations based on DROIs							
ee	31	12.8	9.4	14.8	11.9	34.5	24.1
no	186	13.0	12.1	16.5	14.7	79.0	70.2
se	224	15.8	16.0	18.1	17.9	75.5	63.2
calculations based on CROIs							
ee	14	8.7	5.6	10.4	6.9	25.8	20.3
no	103	8.3	6.9	11.1	9.9	37.4	38.8
se	155	13.0	16.0	14.9	18.1	47.8	82.8

Table A.15: Cross country geometric consistency assessment for **fr**.

CC	#pair$_i$	$\overline{\lvert X_i \rvert}$	$\overline{\lvert Y_i \rvert}$	$\overline{\text{RMSE-}X_i}$	$\overline{\text{RMSE-}Y_i}$	$\lvert \overline{X}_i \rvert_{\max}$	$\lvert \overline{Y}_i \rvert_{\max}$
calculations based on DROIs							
be	134	4.0	5.2	7.5	8.8	21.3	26.1
ch	82	7.8	9.3	10.3	12.4	30.1	33.4
de	214	4.5	6.4	7.0	9.1	18.7	33.2
es	208	3.7	6.2	6.7	9.2	23.3	25.3
gb	5	5.3	10.9	12.6	17.6	10.1	22.0
it	197	9.3	16.6	12.5	19.6	57.1	54.8
lu	39	5.4	4.3	8.0	7.2	22.6	17.3
nl	20	3.3	4.1	9.4	10.4	8.5	17.0
calculations based on CROIs							
be	75	4.7	6.5	9.1	10.6	21.0	32.2
ch	69	7.8	9.5	10.0	12.8	30.4	36.2
de	118	4.5	5.1	6.9	8.7	20.2	23.6
es	120	2.9	6.3	6.0	9.1	15.2	24.9
gb	1	4.6	19.0	10.3	28.2	4.6	19.0
it	96	9.7	18.2	12.4	20.4	32.6	52.6
lu	22	6.4	4.4	8.0	6.1	22.2	13.4
nl	6	2.6	5.6	29.3	28.2	5.8	12.2

Table A.16: Cross country geometric consistency assessment for **gb**.

CC	#pair$_i$	$\overline{\lvert X_i \rvert}$	$\overline{\lvert Y_i \rvert}$	$\overline{\text{RMSE-}X_i}$	$\overline{\text{RMSE-}Y_i}$	$\lvert \overline{X}_i \rvert_{\max}$	$\lvert \overline{Y}_i \rvert_{\max}$
calculations based on DROIs							
fr	5	5.3	10.9	12.6	17.6	10.1	22.0
ie	173	5.9	5.7	10.6	11.2	23.0	33.1
ni	276	5.3	5.0	9.8	10.3	23.3	26.4
calculations based on CROIs							
fr	1	4.6	19.0	10.3	28.2	4.6	19.0
ie	79	6.1	4.8	10.3	10.1	25.4	20.0
ni	265	5.4	5.2	9.9	10.5	23.3	26.4

Table A.17: Cross country geometric consistency assessment for **gr**.

CC	#pair$_i$	$\overline{\lvert X_i \rvert}$	$\overline{\lvert Y_i \rvert}$	$\overline{\text{RMSE-}X_i}$	$\overline{\text{RMSE-}Y_i}$	$\lvert \overline{X}_i \rvert_{\max}$	$\lvert \overline{Y}_i \rvert_{\max}$
calculations based on DROIs							
al	108	13.0	5.3	15.5	7.6	52.9	21.5
bg	176	6.3	7.5	9.2	9.9	27.3	26.5
cs	5	6.0	14.6	7.9	16.6	9.2	19.9
it	2	4.0	6.8	9.3	10.5	6.5	11.8
mc	153	8.7	6.4	11.4	9.5	52.8	27.0
tr	198	11.5	6.4	14.4	9.2	67.8	25.4
calculations based on CROIs							
al	66	7.6	5.0	11.1	8.7	41.7	28.0
bg	81	6.8	5.3	8.6	7.5	27.2	26.2
it	2	3.1	5.6	7.9	10.0	5.3	10.4
mc	79	7.2	5.8	9.8	8.7	25.2	25.8
tr	90	7.1	5.2	10.7	9.3	51.8	19.6

Table A.18: Cross country geometric consistency assessment for **hr**.

CC	#pair$_i$	$\overline{\lvert X_i \rvert}$	$\overline{\lvert Y_i \rvert}$	$\overline{\text{RMSE-}X_i}$	$\overline{\text{RMSE-}Y_i}$	$\lvert \overline{X}_i \rvert_{\max}$	$\lvert \overline{Y}_i \rvert_{\max}$
calculations based on DROIs							
al	4	2.7	5.9	4.2	7.4	5.0	11.5
at	35	5.7	4.7	7.5	7.2	16.7	10.7
ba	207	4.5	5.4	7.8	8.6	28.4	40.6
cs	47	6.0	8.8	12.3	15.2	27.5	53.4
hu	65	7.3	9.4	11.3	13.6	26.9	40.6
it	25	7.1	7.0	9.6	9.8	36.3	21.5
me	22	3.0	3.0	5.4	6.0	7.5	11.4
si	97	5.2	4.5	8.1	7.3	23.2	17.4
calculations based on CROIs							
at	3	21.6	16.7	27.2	21.4	28.2	21.9
ba	151	4.4	5.2	8.5	8.7	27.3	29.1
cs	36	5.2	11.2	11.5	16.0	19.9	53.2
hu	34	7.4	10.6	12.7	15.2	20.4	30.7
it	7	14.2	11.8	26.6	15.8	28.2	17.7
me	10	4.7	2.1	6.1	4.5	13.2	4.8
si	59	3.4	3.6	5.9	5.6	14.6	10.3

Table A.19: Cross country geometric consistency assessment for **hu**.

CC	#pair$_i$	$\overline{\lvert\overline{X}_i\rvert}$	$\overline{\lvert\overline{Y}_i\rvert}$	$\overline{\text{RMSE-}X_i}$	$\overline{\text{RMSE-}Y_i}$	$\lvert\overline{X}_i\rvert_{\max}$	$\lvert\overline{Y}_i\rvert_{\max}$
calculations based on DROIs							
at	84	6.5	6.4	10.0	10.7	24.2	26.1
ba	13	8.7	12.6	16.3	18.0	22.5	32.1
cs	69	7.3	6.2	12.8	11.7	27.7	18.2
cz	14	6.0	7.1	12.9	12.4	15.1	16.3
hr	65	7.3	9.4	11.3	13.6	26.9	40.6
pl	16	4.4	4.6	8.2	8.4	13.2	17.9
ro	81	6.1	7.9	11.8	13.6	24.7	55.2
si	32	7.1	7.7	10.1	11.7	23.4	24.4
sk	157	6.1	5.4	10.9	10.3	23.9	21.3
calculations based on CROIs							
at	43	5.4	7.0	8.7	10.5	16.6	27.0
cs	28	6.7	5.3	10.2	9.6	22.5	13.1
hr	34	7.4	10.6	12.7	15.2	20.4	30.7
pl	3	4.5	4.0	8.1	7.4	9.9	8.4
ro	42	5.0	5.4	11.1	9.6	16.3	20.2
si	18	5.8	11.1	8.4	13.0	17.4	25.9
sk	98	6.1	5.4	11.4	10.5	24.3	23.2

Table A.20: Cross country geometric consistency assessment for **ie**.

CC	#pair$_i$	$\overline{\lvert\overline{X}_i\rvert}$	$\overline{\lvert\overline{Y}_i\rvert}$	$\overline{\text{RMSE-}X_i}$	$\overline{\text{RMSE-}Y_i}$	$\lvert\overline{X}_i\rvert_{\max}$	$\lvert\overline{Y}_i\rvert_{\max}$
calculations based on DROIs							
gb	173	5.9	5.7	10.6	11.2	23.0	33.1
ni	113	5.8	5.6	10.5	11.1	21.4	30.1
calculations based on CROIs							
gb	79	6.1	4.8	10.3	10.1	25.4	20.0
ni	56	6.7	5.2	11.5	11.2	25.4	20.0

Table A.21: Cross country geometric consistency assessment for **it**.

CC	#pair$_i$	$\overline{\lvert\overline{X}_i\rvert}$	$\overline{\lvert\overline{Y}_i\rvert}$	$\overline{\text{RMSE-}X_i}$	$\overline{\text{RMSE-}Y_i}$	$\lvert\overline{X}_i\rvert_{\max}$	$\lvert\overline{Y}_i\rvert_{\max}$
calculations based on DROIs							
al	8	4.3	7.1	9.9	12.3	7.8	14.1
at	123	7.4	5.5	10.2	9.0	31.7	21.8
ch	173	8.8	13.0	13.0	16.3	31.7	41.4
de	51	8.9	10.5	12.4	12.8	26.6	49.5
fr	197	9.3	16.6	12.5	19.6	57.1	54.8
gr	2	4.0	6.8	9.3	10.5	6.5	11.8
hr	25	7.1	7.0	9.6	9.8	36.3	21.5
li	6	10.0	11.7	13.1	13.2	22.0	31.0
mt	1	26.4	67.3	26.4	68.4	26.4	67.3
si	45	6.0	5.2	8.9	7.9	20.2	13.5
calculations based on CROIs							
al	8	3.8	7.0	9.9	13.9	8.1	13.9
at	71	6.2	5.6	8.9	8.6	25.3	23.4
ch	130	6.5	10.2	11.4	13.7	19.3	33.8
fr	96	9.7	18.2	12.4	20.4	32.6	52.6
gr	2	3.1	5.6	7.9	10.0	5.3	10.4
hr	7	14.2	11.8	26.6	15.8	28.2	17.7
si	32	6.3	5.6	8.6	7.5	18.6	15.8

Table A.22: Cross country geometric consistency assessment for **li**.

CC	#pair$_i$	$\overline{\lvert\overline{X}_i\rvert}$	$\overline{\lvert\overline{Y}_i\rvert}$	$\overline{\text{RMSE-}X_i}$	$\overline{\text{RMSE-}Y_i}$	$\lvert\overline{X}_i\rvert_{\max}$	$\lvert\overline{Y}_i\rvert_{\max}$
calculations based on DROIs							
at	11	3.3	1.6	5.9	4.3	11.0	4.6
ch	29	9.9	8.5	11.7	11.2	23.5	34.0
de	23	4.0	8.6	6.6	10.7	16.4	27.1
it	6	10.0	11.7	13.1	13.2	22.0	31.0
calculations based on CROIs							
at	8	3.3	4.1	4.1	5.2	5.8	10.0
ch	10	8.2	9.3	10.1	11.1	17.5	25.5

Table A.23: Cross country geometric consistency assessment for **lt**.

CC	#pair$_i$	$\overline{\lvert\overline{X}_i\rvert}$	$\overline{\lvert\overline{Y}_i\rvert}$	$\overline{\text{RMSE-}X_i}$	$\overline{\text{RMSE-}Y_i}$	$\lvert\overline{X}_i\rvert_{\max}$	$\lvert\overline{Y}_i\rvert_{\max}$
calculations based on DROIs							
ee	3	4.4	7.8	6.9	9.2	9.3	8.4
lv	144	4.2	3.9	6.9	6.7	21.3	20.8
pl	99	7.0	5.1	9.3	7.8	46.9	17.6
calculations based on CROIs							
lv	101	4.7	4.4	7.0	6.9	21.4	21.2
pl	71	6.0	5.6	8.0	7.8	22.3	18.6

Table A.24: Cross country geometric consistency assessment for **lu**.

CC	#pair$_i$	$\overline{\lvert X_i \rvert}$	$\overline{\lvert Y_i \rvert}$	$\overline{\text{RMSE-}X_i}$	$\overline{\text{RMSE-}Y_i}$	$\lvert \overline{X}_i \rvert_{\max}$	$\lvert \overline{Y}_i \rvert_{\max}$
calculations based on DROIs							
be	38	7.5	5.2	9.3	7.7	28.2	18.0
de	42	4.0	5.8	7.0	8.8	18.1	21.4
fr	39	5.4	4.3	8.0	7.2	22.6	17.3
nl	1	7.1	17.8	8.2	19.1	7.1	17.8
calculations based on CROIs							
be	19	10.2	3.2	11.4	5.6	29.8	9.7
de	28	4.2	5.6	6.4	8.3	18.2	17.8
fr	22	6.4	4.4	8.0	6.1	22.2	13.4

Table A.25: Cross country geometric consistency assessment for **lv**.

CC	#pair$_i$	$\overline{\lvert X_i \rvert}$	$\overline{\lvert Y_i \rvert}$	$\overline{\text{RMSE-}X_i}$	$\overline{\text{RMSE-}Y_i}$	$\lvert \overline{X}_i \rvert_{\max}$	$\lvert \overline{Y}_i \rvert_{\max}$
calculations based on DROIs							
ee	68	3.8	3.6	6.4	6.1	19.3	12.2
lt	144	4.2	3.9	6.9	6.7	21.3	20.8
calculations based on CROIs							
ee	42	3.7	3.8	6.9	7.2	19.7	13.2
lt	101	4.7	4.4	7.0	6.9	21.4	21.2

Table A.26: Cross country geometric consistency assessment for **mc**.

CC	#pair$_i$	$\overline{\lvert X_i \rvert}$	$\overline{\lvert Y_i \rvert}$	$\overline{\text{RMSE-}X_i}$	$\overline{\text{RMSE-}Y_i}$	$\lvert \overline{X}_i \rvert_{\max}$	$\lvert \overline{Y}_i \rvert_{\max}$
calculations based on DROIs							
al	60	9.8	4.1	12.9	7.8	46.3	18.5
bg	70	6.4	5.9	8.8	8.2	28.4	23.7
cs	62	6.8	5.2	10.0	8.3	28.4	23.7
gr	153	8.7	6.4	11.4	9.5	52.8	27.0
me	1	4.1	5.1	6.3	6.5	4.1	5.1
calculations based on CROIs							
al	42	8.9	4.8	11.7	8.5	42.5	18.6
bg	41	4.1	3.5	5.7	5.9	12.6	16.7
cs	38	6.8	4.9	9.1	7.5	26.1	22.4
gr	79	7.2	5.8	9.8	8.7	25.2	25.8

Table A.27: Cross country geometric consistency assessment for **me**.

CC	#pair$_i$	$\overline{\lvert X_i \rvert}$	$\overline{\lvert Y_i \rvert}$	$\overline{\text{RMSE-}X_i}$	$\overline{\text{RMSE-}Y_i}$	$\lvert \overline{X}_i \rvert_{\max}$	$\lvert \overline{Y}_i \rvert_{\max}$
calculations based on DROIs							
al	37	4.8	5.7	7.9	8.5	18.1	17.5
ba	48	4.4	5.6	6.6	7.9	20.5	14.9
cs	78	4.4	4.9	6.7	7.3	18.7	14.6
hr	22	3.0	3.0	5.4	6.0	7.5	11.4
mc	1	4.1	5.1	6.3	6.5	4.1	5.1
calculations based on CROIs							
al	26	7.9	4.7	11.3	8.5	19.5	15.8
ba	36	4.1	5.7	6.5	8.2	17.2	18.7
cs	76	4.4	4.8	6.6	7.1	20.9	14.6
hr	10	4.7	2.1	6.1	4.5	13.2	4.8

Table A.28: Cross country geometric consistency assessment for **mt**.

CC	#pair$_i$	$\overline{\lvert X_i \rvert}$	$\overline{\lvert Y_i \rvert}$	$\overline{\text{RMSE-}X_i}$	$\overline{\text{RMSE-}Y_i}$	$\lvert \overline{X}_i \rvert_{\max}$	$\lvert \overline{Y}_i \rvert_{\max}$
calculations based on DROIs							
it	1	26.4	67.3	26.4	68.4	26.4	67.3
calculations based on CROIs							

Table A.29: Cross country geometric consistency assessment for **ni**.

CC	#pair$_i$	$\overline{\lvert X_i \rvert}$	$\overline{\lvert Y_i \rvert}$	$\overline{\text{RMSE-}X_i}$	$\overline{\text{RMSE-}Y_i}$	$\lvert \overline{X}_i \rvert_{\max}$	$\lvert \overline{Y}_i \rvert_{\max}$
calculations based on DROIs							
gb	276	5.3	5.0	9.8	10.3	23.3	26.4
ie	113	5.8	5.6	10.5	11.1	21.4	30.1
calculations based on CROIs							
gb	265	5.4	5.2	9.9	10.5	23.3	26.4
ie	56	6.7	5.2	11.5	11.2	25.4	20.0

Table A.30: Cross country geometric consistency assessment for **nl**.

CC	#pair$_i$	$\overline{\lvert X_i \rvert}$	$\overline{\lvert Y_i \rvert}$	$\overline{\text{RMSE-}X_i}$	$\overline{\text{RMSE-}Y_i}$	$\lvert \overline{X}_i \rvert_{\max}$	$\lvert \overline{Y}_i \rvert_{\max}$
calculations based on DROIs							
be	151	3.0	3.5	6.9	7.7	13.6	25.1
de	205	3.6	4.7	8.6	9.4	19.7	34.5
fr	20	3.3	4.1	9.4	10.4	8.5	17.0
lu	1	7.1	17.8	8.2	19.1	7.1	17.8
calculations based on CROIs							
be	92	3.0	3.3	8.0	8.6	15.2	15.4
de	118	4.1	4.4	11.1	11.4	39.8	21.7
fr	6	2.6	5.6	29.3	28.2	5.8	12.2

Table A.31: Cross country geometric consistency assessment for **no**.

CC	#pair$_i$	$\overline{\lvert X_i \rvert}$	$\overline{\lvert Y_i \rvert}$	$\overline{\text{RMSE-}X_i}$	$\overline{\text{RMSE-}Y_i}$	$\lvert \overline{X}_i \rvert_{\max}$	$\lvert \overline{Y}_i \rvert_{\max}$
calculations based on DROIs							
dk	4	5.2	19.3	14.4	20.7	7.9	22.1
fi	186	13.0	12.1	16.5	14.7	79.0	70.2
se	631	10.8	11.4	14.9	14.3	77.5	53.1
calculations based on CROIs							
fi	103	8.3	6.9	11.1	9.9	37.4	38.8
se	343	9.4	10.7	12.8	13.1	67.7	49.4

Table A.32: Cross country geometric consistency assessment for **pl**.

CC	#pair$_i$	$\overline{\lvert X_i \rvert}$	$\overline{\lvert Y_i \rvert}$	$\overline{\text{RMSE-}X_i}$	$\overline{\text{RMSE-}Y_i}$	$\lvert \overline{X}_i \rvert_{\max}$	$\lvert \overline{Y}_i \rvert_{\max}$
calculations based on DROIs							
at	2	2.6	0.6	4.3	5.1	3.8	0.9
cz	176	4.8	3.7	7.7	6.7	28.3	17.1
de	116	4.1	4.6	8.1	8.2	22.0	22.0
hu	16	4.4	4.6	8.2	8.4	13.2	17.9
lt	99	7.0	5.1	9.3	7.8	46.9	17.6
ro	4	6.5	2.1	11.3	8.3	10.7	5.5
sk	136	5.3	3.1	7.9	6.1	26.9	12.0
calculations based on CROIs							
cz	85	4.2	3.5	6.5	6.1	27.0	20.4
de	60	4.3	4.2	7.7	7.3	26.1	17.7
hu	3	4.5	4.0	8.1	7.4	9.9	8.4
lt	71	6.0	5.6	8.0	7.8	22.3	18.6
ro	2	6.1	2.2	8.7	5.2	8.3	3.5
sk	88	4.9	3.0	6.9	5.5	22.5	12.7

Table A.33: Cross country geometric consistency assessment for **pt**.

CC	#pair$_i$	$\overline{\lvert\overline{X}_i\rvert}$	$\overline{\lvert\overline{Y}_i\rvert}$	$\overline{\text{RMSE-}X_i}$	$\overline{\text{RMSE-}Y_i}$	$\lvert\overline{X}_i\rvert_{\max}$	$\lvert\overline{Y}_i\rvert_{\max}$
calculations based on DROIs							
es	219	4.1	3.3	6.3	5.6	21.7	19.1
calculations based on CROIs							
es	133	4.0	2.9	5.1	4.5	17.8	16.7

Table A.34: Cross country geometric consistency assessment for **ro**.

CC	#pair$_i$	$\overline{\lvert\overline{X}_i\rvert}$	$\overline{\lvert\overline{Y}_i\rvert}$	$\overline{\text{RMSE-}X_i}$	$\overline{\text{RMSE-}Y_i}$	$\lvert\overline{X}_i\rvert_{\max}$	$\lvert\overline{Y}_i\rvert_{\max}$
calculations based on DROIs							
ba	1	0.0	35.9	8.9	37.1	0.0	35.9
bg	194	8.7	8.2	16.5	15.1	54.8	39.9
cs	142	9.0	7.1	13.1	11.4	26.8	50.7
hu	81	6.1	7.9	11.8	13.6	24.7	55.2
pl	4	6.5	2.1	11.3	8.3	10.7	5.5
sk	3	2.9	3.4	7.8	8.5	6.3	6.1
calculations based on CROIs							
bg	97	8.5	8.0	20.8	17.7	50.7	37.6
cs	66	7.4	5.4	12.5	10.9	24.8	19.7
hu	42	5.0	5.4	11.1	9.6	16.3	20.2
pl	2	6.1	2.2	8.7	5.2	8.3	3.5

Table A.35: Cross country geometric consistency assessment for **se**.

CC	#pair$_i$	$\overline{\lvert\overline{X}_i\rvert}$	$\overline{\lvert\overline{Y}_i\rvert}$	$\overline{\text{RMSE-}X_i}$	$\overline{\text{RMSE-}Y_i}$	$\lvert\overline{X}_i\rvert_{\max}$	$\lvert\overline{Y}_i\rvert_{\max}$
calculations based on DROIs							
de	2	5.1	17.6	8.8	21.3	9.3	17.6
dk	68	7.5	9.1	13.8	14.2	48.7	41.8
fi	224	15.8	16.0	18.1	17.9	75.5	63.2
no	631	10.8	11.4	14.9	14.3	77.5	53.1
calculations based on CROIs							
de	1	1.1	12.3	7.7	30.2	1.1	12.3
dk	22	6.2	11.1	14.3	17.1	27.2	26.3
fi	155	13.0	16.0	14.9	18.1	47.8	82.8
no	343	9.4	10.7	12.8	13.1	67.7	49.4

Table A.36: Cross country geometric consistency assessment for **si**.

CC	#pair$_i$	$\overline{\lvert\overline{X}_i\rvert}$	$\overline{\lvert\overline{Y}_i\rvert}$	$\overline{\text{RMSE-}X_i}$	$\overline{\text{RMSE-}Y_i}$	$\lvert\overline{X}_i\rvert_{\max}$	$\lvert\overline{Y}_i\rvert_{\max}$
calculations based on DROIs							
at	113	4.5	5.6	7.0	8.5	17.5	24.2
ba	10	17.5	7.7	20.0	11.1	47.0	21.1
de	1	9.5	12.8	12.5	16.5	9.5	12.8
hr	97	5.2	4.5	8.1	7.3	23.2	17.4
hu	32	7.1	7.7	10.1	11.7	23.4	24.4
it	45	6.0	5.2	8.9	7.9	20.2	13.5
calculations based on CROIs							
at	47	3.9	4.9	5.7	7.0	20.7	24.2
hr	59	3.4	3.6	5.9	5.6	14.6	10.3
hu	18	5.8	11.1	8.4	13.0	17.4	25.9
it	32	6.3	5.6	8.6	7.5	18.6	15.8

Table A.37: Cross country geometric consistency assessment for **sk**.

CC	#pair$_i$	$\overline{\lvert\overline{X}_i\rvert}$	$\overline{\lvert\overline{Y}_i\rvert}$	$\overline{\text{RMSE-}X_i}$	$\overline{\text{RMSE-}Y_i}$	$\lvert\overline{X}_i\rvert_{\max}$	$\lvert\overline{Y}_i\rvert_{\max}$
calculations based on DROIs							
at	48	6.2	4.0	11.1	9.9	23.6	15.7
cz	53	4.4	2.2	8.6	6.8	20.7	8.3
hu	157	6.1	5.4	10.9	10.3	23.9	21.3
pl	136	5.3	3.1	7.9	6.1	26.9	12.0
ro	3	2.9	3.4	7.8	8.5	6.3	6.1
calculations based on CROIs							
at	16	3.2	8.2	6.2	11.4	13.6	25.0
cz	32	5.4	2.3	6.8	4.1	18.8	5.4
hu	98	6.1	5.4	11.4	10.5	24.3	23.2
pl	88	4.9	3.0	6.9	5.5	22.5	12.7

Table A.38: Cross country geometric consistency assessment for **tr**.

CC	#pair$_i$	$\overline{\lvert\overline{X}_i\rvert}$	$\overline{\lvert\overline{Y}_i\rvert}$	$\overline{\text{RMSE-}X_i}$	$\overline{\text{RMSE-}Y_i}$	$\lvert\overline{X}_i\rvert_{\max}$	$\lvert\overline{Y}_i\rvert_{\max}$
calculations based on DROIs							
bg	64	8.1	5.1	11.1	8.3	26.3	14.5
cy	1	23.9	7.2	23.9	8.0	23.9	7.2
gr	198	11.5	6.4	14.4	9.2	67.8	25.4
calculations based on CROIs							
bg	36	6.3	5.7	8.0	7.3	20.4	17.9
gr	90	7.1	5.2	10.7	9.3	51.8	19.6

Bibliography

[1] Adobe Developers Association. TIFF 6.0 specification, 1996. URL http://partners.adobe.com/public/developer/en/tiff/TIFF6.pdf.

[2] A. Annoni, editor. *European Reference Grids*, volume EUR 21494 EN. European Commission, Joint Research Centre, 2005. URL http://www.ec-gis.org/sdi/publist/pdfs/annoni2005eurgrids.pdf.

[3] A. Annoni, C. Luzet, E. Gubler, and J. Ihde, editors. *Map Projections for Europe*, volume EUR 20120 EN. European Commission, Joint Research Centre, 2003. URL http://www.ec-gis.org/sdi/publist/pdfs/annoni-etal2003eur.pdf.

[4] Anonymous. DIMAP Digital Image Map. Technical report, SPOTIMAGE, Toulouse, France, 2008. URL http://www.spotimage.fr/dimap/spec/documentation/refdoc.htm. Last accessed on the 23rd of May 2008.

[5] Anonymous. EPSG Geodetic parameters dataset version.6.15. Technical report, International Association of Oil and Gas Producers, April 2008. URL http://www.ogp.org.uk/.

[6] Anonymous. ISO 3166 code lists: Country names and codes. International Organization for Standardization, 2008. URL http://www.iso.org/iso/country_codes/iso_3166_code_lists.htm.

[7] Anonymous. English country names and code elements. International Organization for Standardization, 2008. URL http://www.iso.org/iso/english_country_names_and_code_elements.

[8] Anonymous. Computation of at-sensor solar exoatmospheric irradiance and Rayleigh optical thickness for IRS-P6 sensors. Technical report, National Remote Sensing Agency, Bangalore, India, unknown. URL http://www.euromap.de/download/exoatm_1.pdf.

[9] Anonymous. Automatic cloud masking of SPOT imagery. *News from Prospace*, (44):42–47, May 1999.

[10] Anonymous. IRS-P6 LGSOWG (Super Structure) Digital Data Products Format. Technical report, Space Applications Centre, ISRO, Ahmedabad, India, May 2003. URL http://www.euromap.de/download/p6super_20050222.pdf.

[11] A. Baraldi, V. Puzzolo, P. Blonda, L. Bruzzone, and C. Tarantino. Automatic spectral rule-based preliminary mapping of calibrated Landsat TM and ETM+ images. *IEEE Transactions on Geoscience and Remote Sensing*, 44(9):2563–2586, September 2006. doi:10.1109/TGRS.2006.874140.

[12] D. Barnea and H. Silverman. A class of algorithms for fast digital registration. *IEEE Transactions on Computers*, C-21:179–186, 1972.

[13] C. Bielski and P. Soille. Order independent image compositing. *Lecture Notes in Computer Science*, 3617:1076–1083, 2005. doi:10.1007/11553595_132.

[14] B. Bishop. Landsat looks at hometown Earth. *National Geographics*, 150(1), July 1976.

[15] V. Dvorchenko. Bounds on (deterministic) correlation functions with applications to registration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 5(2):206–213, March 1983.

[16] H. Fliegel and T. Van Flandern. A machine algorithm for processing calendar dates. *Communications of the ACM*, 11(10):657, October 1968.

[17] D. Hall, G. Riggs, and V. Salomonson. Development of methods for mapping global snow cover using moderate resolution imaging spectroradiometer data. *Remote Sensing of Environment*, 54(2):127–140, November 95. doi:10.1016/0034-4257(95)00137-P.

[18] R. Irish. Landsat 7 automatic cloud cover assessment. In S. Shen and M. Descour, editors, *Proc. of Algorithms for Multispectral, Hyperspectral, and Ultraspectral Imagery VI*, volume 4049, pages 348–355. Society of Photo-Instrumentation Engineers, August 2000. doi:10.1117/12.410358.

[19] R. Irish, J. Barker, S. Goward, and T. Arvidson. Characterization of the Landsat-7 ETM+ automatic cloud cover assessment (ACCA) algorithm. *Photogrammetric Engineering and Remote Sensing*, 72(10):1179–1188, October 2006.

[20] J. Jang, A. Viau, F. Anctil, and E. Bartholomé. Neural network application for cloud detection in SPOT VEGETATION images. *International Journal of Remote Sensing*, 27(4):719–736, Feb 2006. doi:10.1080/01431160500106.

[21] J. Jensen. Urban/suburban land use analysis. In J. Estes, editor, *Manual of Remote Sensing*, volume 2, pages 1571–1666. American Society of Photogrammetry, 1983.

[22] S. Le Hégarat-Mascle and C. André. Reduced false alarm automatic detection of clouds and shadows on SPOT images using simultaneous estimation. In L. Bruzzone, editor, *Proceedings of Image and Signal Processing for Remote Sensing XIII*, volume SPIE-6748, page 674818, Firenze, 2007. doi:10.1117/12.737396.

[23] J. Lewis. Fast normalized cross-correlation. In *Vision Interface*, pages 120–123, 1995.

[24] M. Li, S. Liew, L. Kwoh, and H. Lim. Improved cloud-free multi-scene mosaics of SPOT images. In *Proc. Asian Conf. Remote Sensing*, volume 1, pages 294–298, 1999.

[25] J. Meeus. *Astronomical Algorithms*. Willmann-Bell, Richmond, Virginia, 2nd edition, 1999.

[26] A. Meygret. SPOT Absolute calibration: Synthesis. Technical Report S5-NT-0-2880-CN, CNES, Toulouse, France, August 2007.

[27] R. Müller, T. Krauß, M. Lehner, G. Rönnbäck, and Å Karlsson. Image2006 GMES Fast track land service 2006–2008: Orthorectification of SPOT and IRS-P6 products. Technical report, DLR and Metria, May 2008.

[28] V. Nunes de Lima. *IMAGE2000 and CLC2000: Products and Methods*, volume EUR 21757 EN. European Commission, Joint Research Centre, 2005. URL `http://www.ec-gis.org/sdi/publist/pdfs/nunes2005eur2000.pdf`.

[29] M. Pandya, R. Singh, K. Murali, P. Babu, A. Kirankumar, and V. Dadhwal. Bandpass solar exoatmospheric irradiance and Rayleigh optical thickness of sensors on board Indian Remote Sensing Satellites-1B, -1C,-1D, and P4. *IEEE Transactions on Geoscience and Remote Sensing*, 40(3):714–718, March 2002. doi:`10.1109/TGRS.2002.1000331`.

[30] J. Richards and X. Jia. *Remote sensing digital image analysis*. Springer-Verlag, 3rd edition, 1999.

[31] N. Ritter and M. Ruth. GeoTIFF format specification: Revision 1.0, 2000. URL `http://www.remotesensing.org/geotiff/spec/geotiffhome.html`.

[32] J.-F. Rivest, P. Soille, and S. Beucher. Morphological gradients. *Journal of Electronic Imaging*, 2(4):326–336, October 1993. doi:`10.1117/12.159642`.

[33] P. Soille. *Morphological Image Analysis: Principles and Applications*. Springer-Verlag, Berlin Heidelberg New York, 2nd edition, 2003.

[34] P. Soille. Morphological image compositing. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(5):673–683, May 2006. doi:`10.1109/TPAMI.2006.99`.

[35] P. Soille. IMAGE-2006 Mosaic: Cloud detection on SPOT-4 HRVIR, SPOT-5 HRG, and IRS-LISS III. Technical report, European Commission, Joint Research Centre, 2008.

[36] P. Soille and C. Bielski. Automatic seam line delineation of large image data sets. *IEEE Transactions on Geoscience and Remote Sensing*, 2008. In preparation.

[37] P. Soille and H. Talbot. Directional morphological filtering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(11):1313–1329, November 2001. doi:`10.1109/34.969120`. URL `http://ams.jrc.it/soille/soille-talbot2001r.pdf`.

[38] P. Soille and L. Vincent. Determining watersheds in digital pictures via flooding simulations. In M. Kunt, editor, *Visual Communications and Image Processing'90*, volume 1360, pages 240–250, Bellingham, 1990. Society of Photo-Instrumentation Engineers. doi:`10.1117/12.24211`.

[39] P. Soille, E. Breen, and R. Jones. Recursive implementation of erosions and dilations along discrete lines at arbitrary angles. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(5):562–567, May 1996. doi:`10.1109/34.494646`. URL `http://www.computer.org/tpami/tp1996/i0562abs.htm`.

[40] P. Strobl. Organising a shared raster data repositery for LMNH and SDI units. Technical report, European Commission, Joint Research Centre, 2008.

[41] Q. Tian and M. Huhns. Algorithms for subpixel registration. *Computer Vision, Graphics, and Image Processing*, 35(2):220–233, 1986.

[42] C. Tucker, D. Grant, and J. Dykstra. NASA's Global orthorectified Landsat data set. *Photogrammetric Engineering and Remote Sensing*, 70(3):313–322, March 2004.

[43] L. Vincent and P. Soille. Watersheds in digital spaces: an efficient algorithm based on immersion simulations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(6):583–598, June 1991. doi:10.1109/34.87344.

[44] T. Welch. A technique for high performance data compression. *IEEE Computer Magazine*, 17(6):8–19, June 1984. doi:10.1109/MC.1984.1659158.

[45] P. Wolf. *Elements of Photogrammetry (with Air Photo Interpretation and Remote Sensing)*. McGraw-Hill, New York, 1974.

[46] D. Yuan and C. Elvidge. Comparison of relative radiometric normalization techniques. *ISPRS Journal of Photogrammetry and Remote Sensing*, 51(3):117–126, June 1996. doi:10.1016/0924-2716(96)00018-4.

[47] Y. Zhang, B. Guindon, and J. Cihlar. An image transform to characterize and compensate for spatial variations of Landsat images. *Remote Sensing of Environment*, 82(2–3):173–187, October 2002. doi:10.1016/S0034-4257(02)00034-2.

[48] S. Zucker. Region growing: childhood and adolescence. *Computer Graphics and Image Processing*, 5:382–399, 1976.

European Commission

EUR 23755 EN – Joint Research Centre – Institute for Environment and Sustainability
Title: The IMAGE-2006 Mosaic Project
Editor: Pierre Soille
Luxembourg: Publications Office of the European Union
2011 – 154 pp. – 21.0 x 29.7 cm
EUR – Scientific and Technical Research series – ISSN 1831-9424 (online), 1018-5593 (print)
ISBN 978-92-79-20400-5
doi:10.2788/25572

Abstract

This report details the data ingestion, data organisation, and main processing steps adopted for generating the IMAGE2006 mosaic products. Chapter 1 describes the received data and the way it has been ingested and organised. Chapter 2 presents a detailed analysis of the footprints of the received imagery. Chapter 3 details the procedure used for cloud detection. Relative geometric and radiometric accuracy is studied in Chap. 4. The generation of a cloud mask is indeed fundamental for the generation of a mosaic minimising cloud cover. The mosaicing methodology is presented in Chap 5. The produced mosaics are described in Chap. 6.

The mission of the JRC is to provide customer-driven scientific and technical support for the conception, development, implementation and monitoring of EU policies. As a service of the European Commission, the JRC functions as a reference centre of science and technology for the Union. Close to the policy-making process, it serves the common interest of the Member States, while being independent of special interests, whether private or national.

LB-NA-23755-EN-N

ISBN 978-92-79-20400-5